

# NUTRITIONAL EPIDEMIOLOGY POSSIBILITIES AND LIMITATIONS

NUTRITIONAL  
EPIDEMIOLOGY:  
POSSIBILITIES AND LIMITATIONS

by Lillian Langseth



ILSI Europe

## ILSI EUROPE CONCISE MONOGRAPHS

This booklet is one in a series of ILSI Europe concise monographs. These are written for those with a general background in the life sciences. However, the style and content of the concise monographs will also appeal to a wider audience who seek up-to-date and authoritative reviews on a variety of important topics relating to nutrition, health and food safety.

Concise monographs present an overview of a particular scientific topic and are usually based on the proceedings of scientific meetings. The text of each is peer reviewed by academic scientists of high standing. The concise monographs make important results and conclusions widely available.

Titles of concise monographs in preparation include:

- ◆ Healthy Lifestyles
- ◆ Microwave Ovens.

Current and newly published concise monographs are:

- ◆ Starches and Sugars: A Comparison of Their Metabolism in Man
- ◆ A Simple Guide to Understanding and Applying the Hazard Analysis Critical Control Point Concept
- ◆ Dietary Fibre
- ◆ Food Allergy and Other Adverse Reactions to Food
- ◆ Sweetness – The Biological, Behavioural and Social Aspects
- ◆ Oxidants, Antioxidants, and Disease Prevention
- ◆ Caries Preventive Strategies
- ◆ Nutritional and Health Aspects of Sugars – Evaluation of New Findings
- ◆ Food Biotechnology – An Introduction
- ◆ Dietary Fat – Some Aspects of Nutrition and Health and Product Development
- ◆ Health Issues related to Alcohol Consumption.

## INTERNATIONAL LIFE SCIENCES INSTITUTE AND ILSI EUROPE

The International Life Sciences Institute (ILSI) is a worldwide, public, non-profit scientific foundation headquartered in Washington, DC, USA, with branches in Argentina, Australasia, Brazil, Europe, Japan, Korea, Mexico, North America, Southeast Asia, and Thailand, with a focal point in China.

ILSI is affiliated with the World Health Organization as a non-governmental organization (NGO) and has specialised consultative status with the Food and Agriculture Organization of the United Nations.

ILSI Europe was established in 1986 to provide a neutral forum through which members of industry and experts from academic, medical and public institutions can address topics related to health, nutrition and food safety

throughout Europe in order to advance the understanding and resolution of scientific issues in these areas. ILSI Europe is active in the fields of nutrition, toxicology and food microbiology. It sponsors research, conferences, workshops and publications.

For more information about its programmes and activities, please contact:

ILSI Europe  
Avenue E. Mounier 83, Box 6  
B-1200 BRUSSELS  
Belgium  
Telephone (+32) 2 771 00 14  
Telefax (+32) 2 762 00 44

© 1996 International Life Sciences Institute

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the copyright holder. The International Life Sciences Institute (ILSI) does not claim copyright on U.S. government information.

Authorization to photocopy items for internal or personal use is granted by ILSI for libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Services, provided that \$0.50 per page per copy is paid directly to CCC, 222 Rosewood Drive, Danvers, MA 01923. Telephone: (+1) 508 750 8400.

The use of trade names and commercial sources in this document is for purposes of identification only, and does not imply endorsement by the International Life Sciences Institute (ILSI). In addition, the views expressed herein are those of the individual authors of the original book and/or their organizations, and do not necessarily reflect those of ILSI (or the author of the concise monograph).

ILSI Press

1126 Sixteenth Street, N.W.

Washington, DC 20036-4810

USA

Telephone: (+1) 202 659 0074

Telefax: (+1) 202 659 8654

ILSI Europe

Avenue E. Mounier 83, Box 6

B-1200 Brussels

Belgium

Telephone: (+32) 2 771 00 14

Telefax: (+32) 2 762 00 44

Printed in Belgium

ISBN 0-944398-87-1

Cover figure from Lilienfield DE, Foundations of Epidemiology, © D. Lilienfield 1994 (Used by permission of Oxford University Press)

# FOREWORD

---

Nutritional epidemiology – the study of the nutritional determinants of disease in human populations – is an exciting branch of research because it can provide insight into the potential causes and prevention of many health conditions. But it is also an extremely complex endeavour because many of the associations between dietary factors and disease risks are difficult to discern using epidemiological techniques.

With the increased interest in health-related issues, the results of epidemiological studies easily find their way to the public via newspapers and magazines. The findings, however, must be interpreted with caution. The associations identified by epidemiological studies are important tools for generating and testing new hypotheses about the diet-disease relationship, but other lines of scientific research are needed to provide the evidence that a specific dietary factor causes a health effect.

If applied in appropriate ways, nutritional epidemiology can generate information of great

relevance to public health. This concise monograph aims at making this complex subject understandable. It describes the most common study designs used in nutritional epidemiology and explains their strengths and weaknesses. It also addresses the challenges of accurately measuring food intakes, the use of biomarkers to study nutritional status, the complexities of dealing with bias and confounding factors, the application of statistics and the interpretation of study findings. A series of examples illustrate the possibilities and limitations of nutritional epidemiology.

Public health authorities, health care professionals, scientists in academia and food industry who are not trained in epidemiology, journalists and consumer groups will find this informative publication a useful addition to their libraries. With improved overall understanding of the possibilities and limitations of nutritional epidemiology, people in all of these groups can apply the findings of this field of research in ways that can contribute to current efforts to improve health through disease prevention.

Author: Lillian Langseth

Scientific Editor: Frans J. Kok

Scientific Referees: Pierre Ducimetière, Lenore Kohlmeier, Barrie Margetts, Pirjo Pietinen

Series Editor: Nicholas J. Jardine

# CONTENTS

---

|  |    |
|--|----|
| Introduction .....   | 1  |
| Key features of epidemiology studies: design,exposure assessment and data analysis ..... | 6  |
| Interpretation .....   | 24 |
| Application of nutritional epidemiology data .....                                       | 32 |
| Summary .....  | 35 |
| Glossary .....   | 38 |
| Further reading .....  | 39 |

# INTRODUCTION

---

## Definition of epidemiology

Epidemiology is the study of the distribution and determinants of diseases and other health outcomes in human populations. Epidemiology also deals with the natural history of diseases and it can provide evidence that contributes to their prevention.

Traditionally, epidemiologists devoted most of their attention to infectious diseases. However, even in the early decades of the 20th century, epidemiology made important contributions to the understanding of nutrition-related diseases as well. The investigations of pellagra are a classic example and they illustrate some of the strengths of epidemiology.

### Epidemiological investigation of a deficiency disease

In the 19th and early 20th centuries in both Europe and the United States, pellagra occurred frequently among people who relied on corn as their staple food and who lived in poverty. It was widely suspected that the disease was caused by an infectious agent or by a toxin in spoiled corn. A study by Dr. Joseph Goldberger of the United States Public Health Service, however, suggested otherwise.

Dr. Goldberger observed that pellagra was absent among the staff in an institution even though it was widespread among the residents. Because the staff had access to foods, including meat and milk, that were not served to the residents, Dr. Goldberger began to suspect that the disease might be due to a nutritional deficiency. He obtained funds to supply the residents with meat and milk and found that pellagra rapidly disappeared after these foods were introduced. When the funds ran out and the extra foods were discontinued, pellagra returned.

A few years later, scientists developed an animal model of pellagra and they eventually isolated niacin, the B vitamin that prevents this disease. Three types of scientific investigation – animal experimentation, biochemical studies and epidemiology – all contributed to the eventual full understanding of the causation of pellagra, but it is important to note the key role of epidemiology. At a time when no other type of investigation was possible (because no animal model was available and the biochemical basis of the disease was unknown), epidemiology provided the key clue to solving the mystery of pellagra and epidemiological methods identified a practical means of prevention.

### Epidemiological investigation of chronic diseases

Epidemiology can play a similar role today, even though the current focus is on chronic diseases rather than deficiency syndromes. For example, epidemiological studies in a wide variety of human populations have consistently associated low intakes of fruits and vegetables with increased risks of many types of cancer, as shown in Table 1. The mechanisms by which these foods may protect against cancer are incompletely understood and the active agents have not been conclusively identified. Nevertheless, the epidemiological evidence provides a reasonable basis for action and a sound basis for further research and many authorities now urge people to increase their consumption of fruits and vegetables in an effort to prevent cancer.

## Specific characteristics of nutritional epidemiology

Nutritional epidemiology can be defined as the study of the nutritional determinants of disease in human populations. It is one of the most exciting – and most difficult – types of epidemiological research.

TABLE 1

## Epidemiological studies showing protection by fruits and vegetables against cancer

| Cancer site              | Fraction of studies showing protection |
|--------------------------|--|
| <b>Epithelial</b>        |  |
| Lung                     | 24/25                                  |
| Oral                     | 9/9                                    |
| Larynx                   | 4/4                                    |
| Esophagus                | 15/16                                  |
| Stomach                  | 17/19                                  |
| Pancreas                 | 9/11                                   |
| Cervix                   | 7/8                                    |
| Bladder                  | 3/5                                    |
| Colorectal               | 20/35                                  |
| Miscellaneous            | 6/8                                    |
| <b>Hormone-dependent</b> |  |
| Breast                   | 8/14                                   |
| Ovary/endometrium        | 3/4                                    |
| Prostate                 | 1/14                                   |
| <b>Total</b>             | <b>129/172</b>                         |

Source: Ames BN, Shigenaga MK, Hagen TM, Oxidants, antioxidants and the degenerative diseases of aging, Proc Natl Acad Sci USA 90:7915-7922, 1993.

The excitement comes from the direct relevance of this type of research to crucial health problems of Western civilization, such as coronary heart disease, cancer, stroke, osteoporosis, cataracts, diabetes and congenital malformations. All of these diseases have been the object of research in nutritional epidemiology and some of the findings have already been applied in ways that may improve public health.

For example, epidemiological studies completed in the early 1990s established that women could substantially reduce their risk of bearing a child with a neural tube birth defect (anencephaly or spina bifida) by increasing their intake of the B vitamin folic acid. Even though the mechanism of action of folic acid is not fully understood, public health authorities have begun to take action on this new knowledge. Medical organizations in many nations have recommended increased intakes of folic acid for women of childbearing potential and government agencies in several countries are planning to fortify staple foods with folic acid.

## The complex nature of diet

A major difficulty of nutritional epidemiology lies in the extremely complex nature of diet. To appreciate this complexity, it is helpful to compare diet with another exposure that also influences the risk of many of the same diseases – cigarette smoking.

An epidemiologist who is studying tobacco can obtain a great deal of useful information simply by asking people, "Do you smoke?" By collecting a few additional pieces of information – the number of cigarettes smoked per day, the types or brands of cigarettes smoked, the age at which the person began (or stopped) smoking and any changes that may have occurred in the pattern of cigarette use – the researcher can obtain a clear, reasonably accurate picture of an individual's smoking history.

In contrast, one cannot learn much by asking people, "Do you eat?" Everyone eats. Everyone is exposed, to varying degrees, to most dietary factors. Eating patterns often evolve slowly over periods of years and people may not remember when their habits changed. The foods that people eat consist of complex mixtures of compounds, with substantial differences even among seemingly similar products. People who eat more of one type of food must eat less of other types of foods, thus creating a complex set of intercorrelations among dietary components. Eating habits may be correlated with other factors that influence disease risk, such as ethnic background, socioeconomic status and tobacco use. Even the method of preparation of foods may be important. For example, boiled coffee may raise blood cholesterol levels; filtered coffee does not because filtering removes the components that may have cholesterol-raising effects.

#### Multifactorial causation of chronic diseases

Like diet, the diseases currently of interest to epidemiologists are also complex. They have multiple determinants and long latent periods and they occur with relatively low frequencies, even among people with high exposure to risk factors.

For example, coronary heart disease has a wide variety of recognized risk factors including age, gender, menopausal status in women, family history, body weight, blood pressure, blood cholesterol and diabetes. Other factors, such as the degree of oxidation of blood lipoproteins and levels of the amino acid homocysteine, may also be involved. Many of these risk factors and suspected risk factors are influenced by multiple aspects of diet. For example, intakes of several types of fatty acids influence blood cholesterol levels and intakes of three different B vitamins influence homocysteine levels. Some risk factors exert their effects over long periods of time (usually by influencing the progression of atherosclerosis). Others, however, may exert their

effects very quickly (by influencing the likelihood of blood clotting). The more that scientists learn about the nature of coronary heart disease, the more complex the causes seem to be.

In light of this complexity, it is easy to see why scientists do not fully understand the reasons for differences in coronary heart disease rates at different times and places. Even dramatic effects – such as the unexpectedly low rate of coronary disease in France or the decrease of more than one-third in cardiovascular mortality in the United States in recent decades – cannot be easily explained.

#### The inherent limits of epidemiology

In recent years, it has become increasingly evident that the science of epidemiology has inherent limits. Although epidemiology is very effective in identifying strong links between an environmental factor and a disease (for example, the link between smoking and lung cancer), it is less effective in discerning weaker associations. As Dr. Michael Thun of the American Cancer Society stated in a recent interview in the journal *Science*, "With epidemiology you can tell a little thing from a big thing. What's very hard to do is to tell a little thing from nothing at all." Many of the associations between diet and disease are "little things": if the effects are real, they are relatively subtle. It may be impossible to determine, from epidemiology alone, whether relatively weak associations between diet and disease are real or whether they reflect some type of subtle bias or measurement error that the researchers were unable to eliminate.

## Goals of nutritional epidemiology

Nutritional epidemiology has several goals. Perhaps the most basic is monitoring the food consumption, nutrient intake and nutritional status of a population. Other key goals are to generate new hypotheses about diet and disease, to produce evidence that supports or refutes existing hypotheses and to assess the strength of diet-disease associations. Ultimately, the overall goal of nutritional epidemiology is to contribute to the prevention of disease and the improvement of public health.

Epidemiology alone cannot reach this goal, however. Most types of epidemiological studies (with the notable exception of intervention trials, as described in a later section) can only identify associations; they cannot prove that an exposure causes a health effect. When intervention trials are not possible, other scientific methods must be combined with epidemiology to establish a convincing causal relationship.

## Advantages and disadvantages of nutritional epidemiology

### Advantages

As several of the previously cited examples illustrate, a key advantage of nutritional epidemiology is its direct relevance to human health. Epidemiologists study real life. They do not need to extrapolate from animal models or *in vitro* systems. The results of their work are often used to calculate direct estimates of risk, which can then be translated into specific recommendations for changes in nutrient intakes or food consumption patterns.

Findings from nutritional epidemiology can even have direct implications for food processing and technology. For example, recent epidemiological studies associating high intakes of trans fatty acids (found in margarine and other processed vegetable fats) with increased risks of

coronary heart disease will probably prompt margarine manufacturers to seek out ways to reformulate their products to reduce their trans fatty acid content.

### Disadvantages

Perhaps the most important disadvantage of nutritional epidemiology is the potential for many kinds of bias. Bias is defined as systematic error, resulting in over- or underestimation of the strength of an association between an exposure and an outcome. As discussed in more detail in a later section, studies in nutritional epidemiology must be designed and executed with great care to minimize bias.

Another disadvantage of epidemiology is the difficulty in determining whether observed associations are causal. If the association between a factor and a disease is not causal, efforts to modify exposure to that factor will not reduce disease risk. For example, even though the drinking of alcohol is associated with lung cancer risk, efforts to discourage alcohol consumption would not be likely to reduce the lung cancer death rate, because the relationship is not causal. Instead, it reflects the association of both alcohol intake and lung cancer with a third factor – cigarette smoking.

The apparent simplicity and “real life” relevance of epidemiological findings may also be a disadvantage in some circumstances because they encourage the misuse and overinterpretation of data. This is especially true when preliminary or unconfirmed findings come to the attention of the news media and the general public. For example, the reports of an association between margarine intake and cardiovascular disease may have prompted some consumers to switch back to butter, even though most experts believe that this course of action would not be beneficial to cardiovascular health.

## BOX 1

### Nutritional epidemiology in action: alcohol and breast cancer

Breast cancer is one of the most common types of cancer among women in the industrialized countries. Efforts are being made to identify modifiable risk factors for this disease in the hope of finding ways to reduce its incidence. One possible risk factor for breast cancer that has been singled out by epidemiological studies is alcohol intake. If any association exists between alcohol intake and breast cancer risk, it could be of huge importance to public health because large numbers of women drink alcoholic beverages.

More than 50 epidemiological studies have examined the relationship between alcohol intake and breast cancer. More than half of these studies have reported that drinkers have an increased risk of the disease. Typically, the relative risk in women who drink alcohol, as opposed to nondrinkers, is about 1.5.

Is this a true relationship? No one really knows. The relative consistency of the evidence tends to argue in favour of a true association. However, the increase in risk is so small that even a subtle bias might account for it. Moreover, no mechanism by which alcohol might influence breast cancer risk has been established. Some scientists argue strongly in favour of a link between alcohol and breast cancer, whereas others argue with equal fervour that no such association has been convincingly demonstrated.

It may be impossible to confirm or refute a link between alcohol and breast cancer by epidemiological means. The effect is so weak that observational studies cannot establish with certainty whether it is real. Indeed, this is an excellent example of the kind of situation that pushes observational epidemiology to its inherent limits and forces scientists to be very cautious in interpreting such studies. If an intervention trial were possible, it might settle the question, but such a study is unlikely to take place. It would be extremely difficult (and perhaps unethical) to randomly assign large numbers of women to drink or abstain from alcohol for years on end. In the absence of intervention trials, a definitive answer to the alcohol–breast cancer question must await new findings from other types of research – such as animal experiments or biochemical studies. Epidemiology alone is unlikely to resolve this issue.

# KEY FEATURES OF EPIDEMIOLOGY STUDIES: DESIGN, EXPOSURE ASSESSMENT AND DATA ANALYSIS

---

## Study design

Several different study designs are used in nutritional epidemiology. These include descriptive studies, ecological studies, case-control studies, cohort studies and intervention trials. Each study design serves different purposes and has specific strengths and weaknesses.

It is important to distinguish between study designs that employ aggregate data on populations and those that use individual data on specific members of a population. As a general rule, individual-based studies are of greater value than those using aggregate data and their results should be given greater weight.

## Descriptive epidemiology

Epidemiology studies can be classified into two categories: descriptive and analytic. Descriptive epidemiology is the study of the amount and distribution of diseases, exposures or other factors of interest within a population in terms of person, place and time. Analytic epidemiology is the more specific study of the determinants of disease in study populations.

In descriptive studies, researchers may collect data on a variety of factors including dietary intakes, biochemical markers of nutritional status, risk factors for disease, the incidence of disease or rates of death from various causes. They may then compare the findings for different types of persons: those of different ages, genders, ethnic backgrounds and socioeconomic or

occupational groups. They may compare the data obtained from people who live in different places, that is, different geographic areas or different types of communities (for example, urban versus rural). If repeated studies are conducted, information may also be generated on changes in nutritional variables over time.

The data generated by descriptive epidemiology can be used to identify variations in the distribution of disease or in the distribution of nutritional factors. The findings are also useful for comparing actual dietary intakes with established standards such as Recommended Dietary Allowances (RDAs). Many descriptive data are collected for the purpose of nutritional surveillance, which is beyond the scope of this monograph.

## National surveys

In some countries, periodic national surveys provide extensive descriptive data that can be of great value to nutritional researchers. For example, in the United Kingdom, scientists can make use of the annual National Food Survey (NFS), which records food purchases over a 1-week period in a representative sample of 7 500 British households. The reports from the NFS provide an excellent unbroken record of British food habits since 1950.

In the United States, researchers can investigate a broad range of topics using the extensive data collected in the three successive National Health and Nutrition Examination Surveys (NHANES I, II and III). For example, one United States research group recently used data from NHANES I to investigate whether users of vitamin supplements have lower death rates (they do not), a second group used NHANES II data to find out whether Americans are eating the recommended number of daily servings of fruits and vegetables (they do not even come close) and a third group compared data from the three successive NHANES surveys to see whether Americans are reducing their fat intake (they are, but not as rapidly as public health authorities would like).

## Ecological epidemiology

Studies using aggregate data on diet and disease, also called ecological studies, compare indices of nutritional status with indices of health status for groups of people. This type of research is usually conducted in the early stages of the investigation of a diet-disease relationship. Because ecological studies are relatively quick and inexpensive, often relying on data that have been routinely collected for other purposes, it makes sense to use them for a preliminary evaluation of a new hypothesis to determine whether more extensive and costly investigations are warranted.

### Advantages of ecological studies

In some situations, ecological studies may be the only way to test a hypothesis. Data on a particular exposure or outcome may be available only for communities, not for individuals. For example, it is generally not feasible to measure individual exposure to nutrients or toxic substances in drinking water. Researchers can, however, easily determine the amounts of these substances in the water supplies of different communities and compare them with health outcomes.

By the late 1930s, it was recognized that the use of water supplies with high fluoride concentrations led to mottling of tooth enamel. Dentists observed that people with mottled teeth had low rates of dental caries and they hypothesized that fluoride might prevent caries. The United States Public Health Service tested this hypothesis by conducting an ecological study. Researchers surveyed the dental health of children in 13 cities where the fluoride concentration in the water supply varied considerably. The results, reported in 1942, indicated that dental caries decreased with increasing fluoride content of water. A fluoride level of one part per million appeared to be optimal; it was associated with a reduction in caries without an unsightly degree of mottling. These ecological findings set the stage for later experimental studies that

established the benefit of adjusting the fluoride content of drinking water supplies to one part per million.

Ecological studies may also be the best type of investigation when the variation in exposure to a dietary factor within a population is so small that differences in health outcomes cannot be detected. Some scientists have argued, for example, that the relationship between dietary fat intake and breast cancer risk is best addressed through ecological comparisons of different populations, since the fat intakes of individuals within a single population tend to be relatively homogeneous.

### Limitations of ecological studies

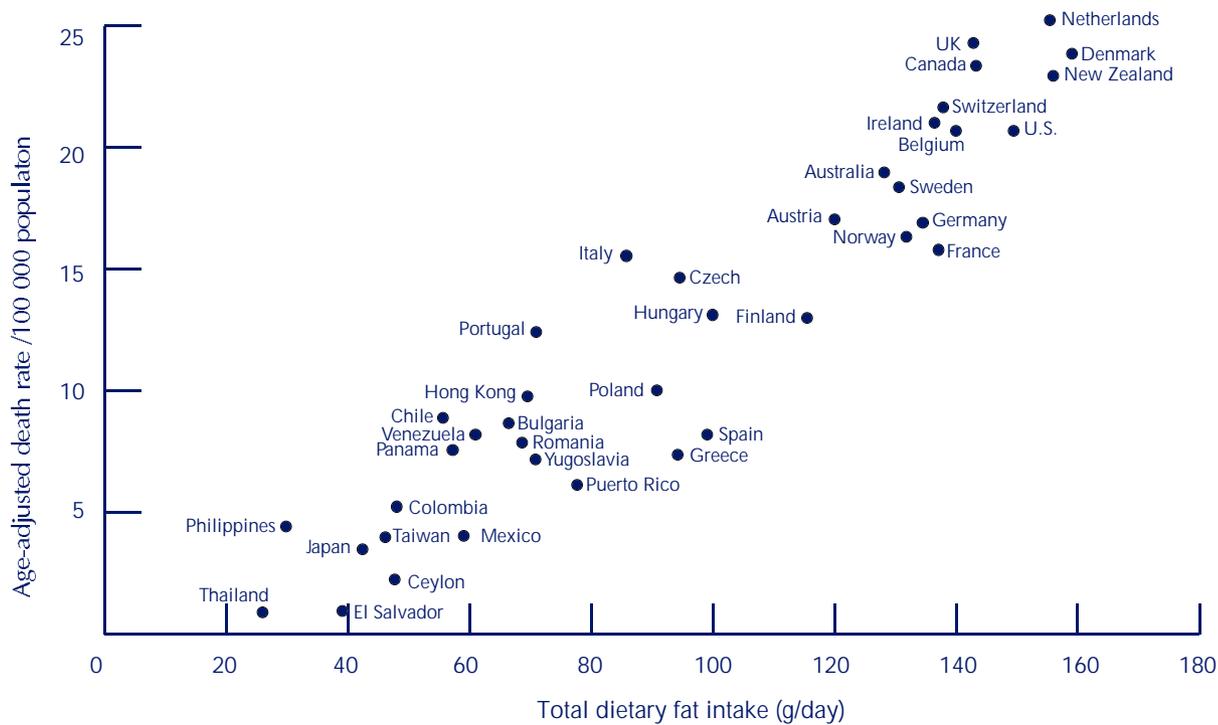
A key limitation of ecological studies is that findings for groups of people may not be applicable to individuals. Even if aggregate data show that populations with higher fat intakes have higher rates of coronary disease, it does not necessarily follow that the individuals who have heart attacks are the higher consumers of fat. Nor does it follow that differences in fat intake were necessarily responsible for the differences in coronary disease rates; the populations may have differed in other ways that could be relevant to cardiovascular health.

### Types of ecological studies

One common type of ecological study, often referred to as a cross-cultural correlation study, involves the comparison of food consumption data and disease rates in different countries or geographic areas. Studies of this type played an important role in the early stages of the investigation of the relationship between diet and cancer. More than 20 years ago, cross-cultural studies revealed very strong positive correlations between meat and fat consumption in various countries and the rates of breast and colon cancer in those countries. Figure 1 shows a classic set of data on fat intake and breast cancer. Findings such as these prompted extensive further investigations of the possible link between dietary fat and cancer.

FIGURE 1

Relation of national per capita fat intake to risk of breast cancer



Source: Willett W, Nutritional Epidemiology. (© Walter S. Willett 1989 (Used by permission of Oxford University Press)).

A major problem with cross-cultural correlations, however, is that many aspects of a population's lifestyle may be correlated. For example, intakes of fat and meat are highly correlated with economic development. Other measures associated with development, such as gross national product, also show strong correlations with breast and colon cancer rates. Thus, it is unclear from ecological studies whether intakes of fat and meat have a true impact on cancer risk or whether they merely serve as markers for other factors that change with economic development, such as physical activity. Other types of epidemiological studies have not consistently confirmed the strong associations between dietary fat intake and breast and colon cancers seen in cross-cultural comparisons. Despite decades of investigation, it remains uncertain whether there is a causal relationship between dietary fat intake and the risk of cancers at these body sites.

Ecological studies may also be carried out within a country if there are substantial differences in eating habits in different regions or subpopulations. For example, researchers in Belgium compared nutritional and health indices of the Flemish-speaking and French-speaking populations within their country and found that those who spoke French had higher dietary fat intakes, higher serum cholesterol levels and higher rates of coronary heart disease.

Ecological studies using individual exposure data Occasionally, ecological studies may be carried out using data on individual members of several populations. A classic example is the Seven Countries Study, in which individual-based studies of coronary heart disease were carried out in 16 populations in seven countries. The data from these studies were compared in an effort to identify cross-cultural differences.

### Time trends

Some ecological studies compare populations from different times rather than different places. For example, after animal experiments suggested that saccharin might cause bladder cancer, trends in bladder cancer rates were analyzed to see whether they could be linked with changes in the use of this sweetener. (No such relationship was found.)

### Special populations

Other types of ecological studies involve special populations that are particularly relevant to the disease or dietary factor under investigation. For example, some epidemiological studies of saccharin and bladder cancer focused on people with diabetes, because they are especially heavy users of artificial sweeteners. (Again, no link was found between saccharin and bladder cancer.)

Other special populations of interest to epidemiologists include religious groups such as Seventh-Day Adventists, who follow distinctive lifestyles. It is important to remember, however, that these groups may differ from the general population in multiple ways. For example, many Adventists are lacto-ovo-vegetarians and it would be tempting to conclude that this dietary pattern is responsible for their reduced risks of several chronic diseases. However, Seventh-Day Adventists also abstain from the use of alcohol and tobacco and these practices may be largely responsible for their distinctive patterns of disease.

### Migrant studies

Populations that have migrated from one part of the world to another provide a unique opportunity for assessing the relative contributions of genetics and environmental factors to disease. Diet, of course, is one of the environmental factors that change when people move from one place to another and gradually adopt the lifestyle of their new country or community.

## 10 Concise Monograph Series

Extensive studies have been conducted on several migrant populations including Japanese migrants to the United States and European migrants to Australia. These studies have shown that migrant groups eventually acquire the disease patterns of their new countries but that the change occurs more rapidly for some diseases than for others.

### Some epidemiological terminology

Incidence refers to the number of new cases of a disease in a specific time period. It is usually expressed as an incidence rate: the number of new cases over a specified period of time divided by the number of people in the population at risk of developing the disease. The use of a rate enables populations of different sizes to be compared directly.

Incidence data are very useful because they provide a direct measure of the rate at which individuals in a population develop a disease and therefore provide a basis for statements about the probability or risk of disease. However, differences in incidence can be difficult to interpret because they may reflect patterns in the diagnosis of a disease rather than its true occurrence. For instance, the 46% increase in the incidence of prostate cancer in the United States between 1973 and 1987 was attributable primarily to increased screening efforts and the availability of sensitive new diagnostic tests, rather than to changes in diet or other risk factors.

Prevalence refers to the number of existing cases of a disease or other condition in a population. For example, in the initial examination of participants in the Framingham Study (a long-term study of coronary heart disease in a United States community), 43 of the 941 men 45–62 years old had coronary heart disease, for a prevalence of 46 per 1 000.

It is important to note that the prevalence of a disease

may reflect changes in the length of survival of the patients (and therefore possibly changes in treatment) as well as changes in incidence. The prevalence of diabetes mellitus no doubt increased dramatically after the discovery of insulin, since patients no longer died within a short time after diagnosis. There is no particular reason to suspect, however, that the incidence of the disease changed at that time.

Nutritional epidemiologists are interested in the prevalence of risk factors for disease as well as in the prevalence of disease itself. Data on the prevalence of risk factors, such as overweight and smoking, can help authorities set priorities for interventions. Comparison of the data from the successive NHANES surveys has shown that the prevalence of overweight among United States adults increased from 25% to 33% in one decade. This finding clearly indicates the need for increased attention to this major health risk factor.

Mortality (death) rate data are sometimes easier to obtain than incidence or prevalence data, but they are greatly influenced by changes in the treatment of disease and do not simply reflect its occurrence. For some diseases of interest to nutritional epidemiologists, such as cancers of the lung and pancreas, mortality rates are very similar to incidence rates because the diseases are rapidly and almost uniformly fatal. For other diseases, however, incidence and mortality patterns are very different. Prostate cancer is a good example. The 46% increase in the incidence of this disease in the United States was accompanied by only a 7% increase in mortality from this cause.

For some diseases, mortality data are of no value because deaths are rare, even though the burden of disease may be very great. People do not die from cataracts, for instance, but the treatment of this condition imposes a large burden on the health care budgets of Western countries. In the United States alone,

physicians perform 1.2 million cataract extractions annually, at a total cost of over \$3.2 billion. If ways could be found to prevent cataract, perhaps through better diets, the benefits would be great. No change in mortality would be expected, however.

#### The need for age-standardization

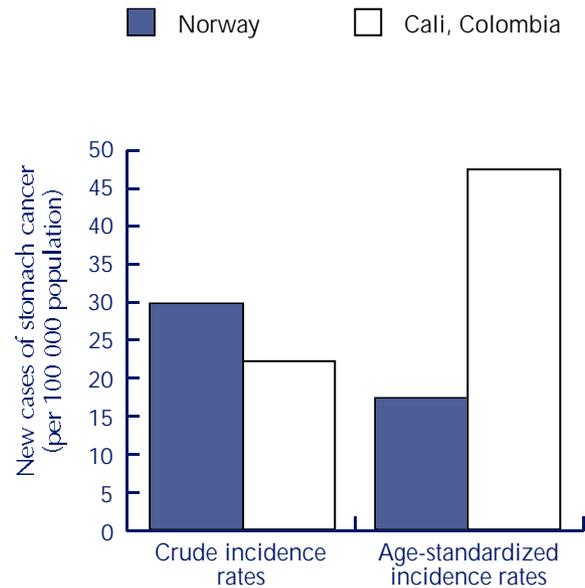
Most diseases do not occur with equal frequency among people of all ages. Many of the diseases of greatest interest to nutritional epidemiologists, including coronary heart disease, stroke and cancer, become dramatically more frequent with increasing age. For this reason alone, the rates of these diseases are higher in populations that include a large proportion of older adults (such as the current populations of Western Europe and North America) than in those that include fewer elderly people (such as the current populations of developing nations or the past populations of industrialized countries).

For a meaningful comparison of disease rates in populations with different age structures, it is necessary to take account of the effects of age. This can be done by comparing age-specific rates (for example, the incidence of stroke among women 60–64 years old) or by using a technique called age standardization, in which the rates of disease in each age group are weighted by the age distribution of a standard population. The resulting rates can then be compared in a meaningful way.

To illustrate, consider the incidence rates of stomach cancer among men in Norway and men in Cali, Colombia, as shown in Figure 2. For all age groups, the age-specific incidence of stomach cancer in Cali is higher than that in Norway. However, because Norway has an older population (18% of men in Norway are 60 years old or older, compared with 4.6% in Cali), the crude incidence rate of stomach cancer in Norway is actually slightly higher than that in Cali (29.7 per 100 000 versus 23.2 per 100 000). When the stomach

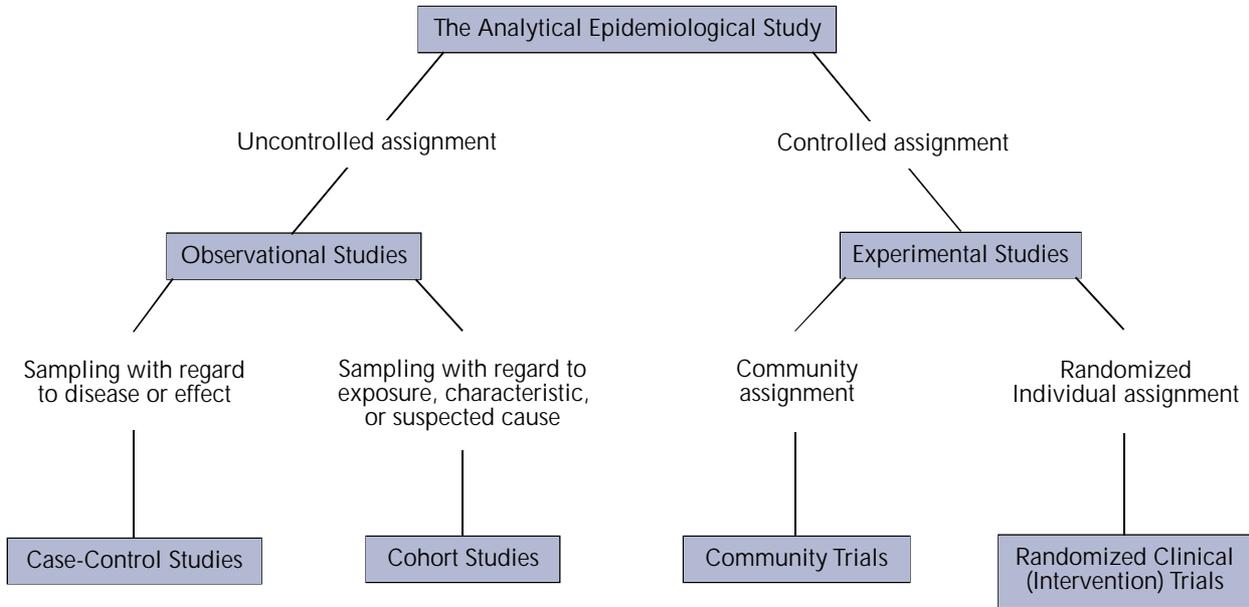
FIGURE 2

Comparison of crude and age-standardized stomach cancer rates in two populations



Source: Margetts BM, Nelson, M (eds), Design Concepts in Nutritional Epidemiology, (© Oxford: Oxford University Press, 1991).

FIGURE 3  
Types of epidemiological studies



Source: Lilienfeld DE, Foundations of Epidemiology, © D. Lilienfeld 1994 (Used by permission of Oxford University Press)

cancer incidence rates for men in Norway and Cali are standardized by comparison to the world standard population, a different and more meaningful picture emerges. After standardization, the rate for Norway is 18.1 per 100 000 and that for Cali is 48.3 per 100 000. These rates accurately reflect the age-specific rates and correctly suggest that men in Cali are exposed to more risk factors for stomach cancer than are men in Norway.

#### Studies using individual data

For purposes of hypothesis testing, studies using individual data on diet and disease are generally preferred. Figure 3 shows the principal types of studies.

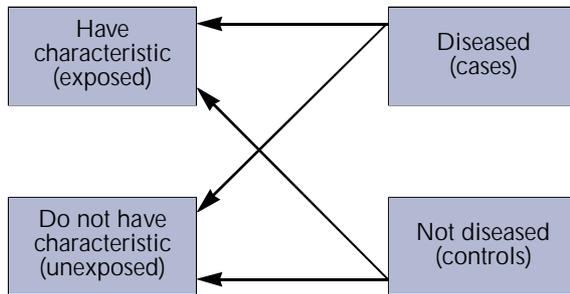
Individual-based studies may be experimental (exposures are manipulated by the investigator) or observational (they are not). The two most important observational study designs are the case-control study and the cohort study.

#### Case-control studies

For a case-control study, researchers identify people who have a disease (cases) and otherwise similar people who do not have it (controls) and compare their exposures to factors that may have influenced their disease risk. Ideally, the cases and controls should be selected from the same population and they should be representative of that population. Case-control studies

FIGURE 4

The basic principle of case-control studies



Source: Ahlbom A, Norell S, Introduction to Modern Epidemiology, (© Newton Lower Falls, MA: Epidemiology Resources, Inc., 1984).

are retrospective, meaning that the focus is on exposures that occurred in the past and on the ways in which these exposures may have affected an individual's present health. Figure 4 illustrates the basic principle of a case-control study.

For example, a case-control study of chilli pepper consumption and stomach cancer was conducted in Mexico City. A total of 220 patients with stomach cancer (cases) and 752 stomach cancer-free but otherwise similar people selected from the general population (controls) were interviewed about their chilli pepper consumption and other aspects of their diets. It was found that a higher proportion of cases than controls ate chilli peppers. To state it in another way, the likelihood of stomach cancer in chilli pepper consumers appeared to be greater than that in non-consumers.

**Advantages of case-control studies.** Case-control studies have many advantages. They can investigate a wide variety of potential risk factors simultaneously, they are

relatively quick and inexpensive (in comparison to cohort studies, as described below) and they can be applied a priori to all diseases, both common and rare. Another advantage of this type of study is that it is possible to match subjects for important factors that are not currently under investigation. For example, in a case-control study of lung cancer, the risk of which is profoundly influenced by cigarette smoking, one could select controls with histories of smoking as similar as possible to those of the cases, so that attention could be focused on other factors such as diet. (Even if matching is not performed, it is important to collect information on factors that may influence risk but are not currently under investigation, so that appropriate adjustment for these factors can be made during data analysis, as discussed in a later section)

**Disadvantages of case-control studies.** Case-control studies also have major disadvantages. This type of investigation requires the researcher to collect information about the subjects' past exposures – a difficult task. People's memories of past dietary habits are often faulty and objective data on past exposures (for example, biological markers) may not be available.

**Types of bias.** Case-control studies are also subject to several types of bias, including selection bias and information bias. Selection bias occurs when study subjects are chosen in a way that can misleadingly increase or decrease the magnitude of an association. This can occur when the cases and controls are selected from different populations or when the subjects in either group are not representative of the population from which they come.

Researchers often select their control subjects from among the other patients in the same hospitals where the cases were treated. The use of hospital controls is convenient, but it may not be appropriate for some types of studies, such as those involving alcohol intake. Data indicate that hospital patients drink more alcohol

## 14 Concise Monograph Series

than the general public, so their hospitalization may reflect their increased risk of alcohol-related illnesses and accidents. This in turn may lead to bias in studies where alcohol consumption is an important exposure variable. According to a recent review, nine out of ten community-based studies of alcohol intake and colorectal cancer risk showed a positive association between these two factors, but only five out of 17 hospital-based studies showed a similar association. The negative results of the hospital-based studies may have been due to bias.

To avoid problems of this sort, researchers may choose to use control subjects selected from the general population. Unfortunately, however, bias is also possible in the selection of these controls. If researchers search for controls by dialing randomly chosen telephone numbers, for example, the control group will not include people who do not have telephones. Because people who lack telephones tend to be of low socioeconomic status, and because socioeconomic status influences the risk of disease, the absence of these people from the control group may bias the study's findings.

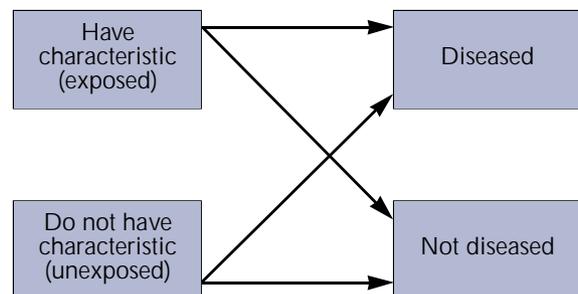
Information bias results when the method of data collection makes the information obtained from cases and controls different in a misleading way. For example, cases may recall past events differently than healthy controls do because they are motivated to pay more attention to the causes of disease; this is called recall bias. If the interviewer knows whether subjects are cases or controls (for example, because cases are visibly ill), the conduct of the interview may change in subtle ways, leading to interviewer bias. Epidemiologists who conduct case-control studies need to plan their research so that both recall bias and interviewer bias are reduced as much as possible.

Information bias may also occur when biological markers are used as an index of nutritional status. The levels of some markers in the cases may be modified by the onset of disease (for example, fatty acids after a myocardial infarction).

Another type of bias encountered in epidemiological studies is regression dilution bias, in which the variability in a biological measurement may lead to an underestimation of the strength of an association. For example, blood cholesterol levels vary from day to day within a single individual. If researchers measure their subjects' cholesterol levels on a single occasion and then assess the association between these levels and the risk of heart disease, the observed association is likely to be weaker than the true association because of the variability in cholesterol measurements. This problem can be dealt with by making repeated measurements of parameters that tend to fluctuate or by using mathematical methods to account for the impact of dilution bias.

FIGURE 5

The basic principle of cohort studies



Source: Ahlbom A, Norell S, Introduction to Modern Epidemiology, (© Newton Lower Falls, MA: Epidemiology Resources, Inc., 1984).

### Cohort studies

Unlike a case-control study, which begins with a group of people who already have a disease, a cohort study begins with a defined study population that is followed over a period of time. Cohort studies may be either prospective or retrospective. In a prospective cohort study, the study population is characterized at the start of the study and followed into the future. In a retrospective or historical cohort study, a cohort that was characterized in the past is followed to the present. The retrospective cohort study design is rarely used in nutritional epidemiology and will not be discussed further in this monograph.

For a prospective cohort study, researchers identify a population of people who do not (yet) have the disease under investigation and collect information on the subjects' exposure to risk factors, including nutritional factors. They then follow the study subjects for a period of time to see who develops the disease. The frequency of disease among subjects exposed to a particular risk factor is compared to the frequency among those who were not exposed. Figure 5 shows the basic principle of a prospective cohort study.

As part of the Zutphen study, a prospective cohort study of cardiovascular disease conducted in the Netherlands, 552 men provided information in 1970 on their current dietary habits, including their consumption of fish. During the next 15 years, 42 of these men suffered strokes. The researchers discovered that fish intake in 1970 was lower among men who later experienced strokes than among those who did not.

**Advantages of cohort studies.** A major advantage of a cohort study is its prospective design, which avoids many of the problems common to retrospective case-control studies. Subjects are not asked to recall past events and bias is minimized because the data are collected before the development of disease.

**Disadvantages of cohort studies.** Unlike the subjects in a case-control study, the participants in a cohort study must be contacted repeatedly to determine their health outcomes. This requires long-term record keeping and some degree of cooperation from the subjects. To facilitate this, scientists sometimes select their cohorts from among people who can be expected to cooperate well, such as health professionals. However, such individuals may not be representative of the general population.

Because researchers must wait for health outcomes to occur and because many of the diseases they are studying take many years to develop, cohort studies are often very long in duration. Because disease frequencies are low, even for common diseases, it may be necessary to enroll large numbers of subjects (at great expense) to obtain meaningful results. Because scientists often do not know how long it might take for an exposure to modify the risk of disease, it may be necessary to determine the outcomes at several points in time. If the researchers wish to assess the effects of changes in exposure, the exposures may also need to be measured at several points in time, thus further increasing the cost and complexity of the study. When large numbers of subjects are used, cost considerations usually dictate that exposures can be measured only in relatively crude ways (for example, by mailed questionnaires) rather than by the extensive personal interviews, physical examinations and laboratory tests that could be used in a case-control study.

### Nested case-control studies

Occasionally, scientists can combine some of the advantages of a cohort study with those of a case-control study by conducting a nested case-control study, in which the cases and controls are selected from a cohort about which information was collected on a prior occasion. Studies of this type can combine the speed and low cost of a case-control study with the relative freedom from bias of a cohort study.

## BOX 2

## Nutritional epidemiology in action: vitamin E supplements and heart disease

Two cohort studies in the United States have associated the use of high-dose vitamin E supplements with substantial reductions in the risk of coronary heart disease. In one of these studies, which involved more than 87 000 women, the use of vitamin E supplements for two or more years was associated with a 41% decrease in coronary risk. In the parallel study of almost 40 000 men, vitamin E supplementation was associated with a 37% reduction in risk. In both studies, the types of supplements associated with risk reduction were those that provided relatively high doses of vitamin E – 100 mg or more per day. Multivitamins and dietary vitamin E had no significant effect.

A protective effect of vitamin E against heart disease is biologically plausible. Vitamin E is an antioxidant. As such, it may protect against harmful oxidative processes that play a role in the causation of disease. One such oxidative process – the oxidation of low-density lipoproteins (LDL) – is believed to contribute to the causation of atherosclerosis. Studies in human subjects have shown that oral supplementation with large doses of vitamin E decreases the susceptibility of LDL to oxidation. A dose-response study has shown that this effect is observed only at doses of at least several hundred milligrams of vitamin E per day. This dose-response relationship is consistent with the two epidemiological studies in which significant associations were found only for relatively high dose supplements.

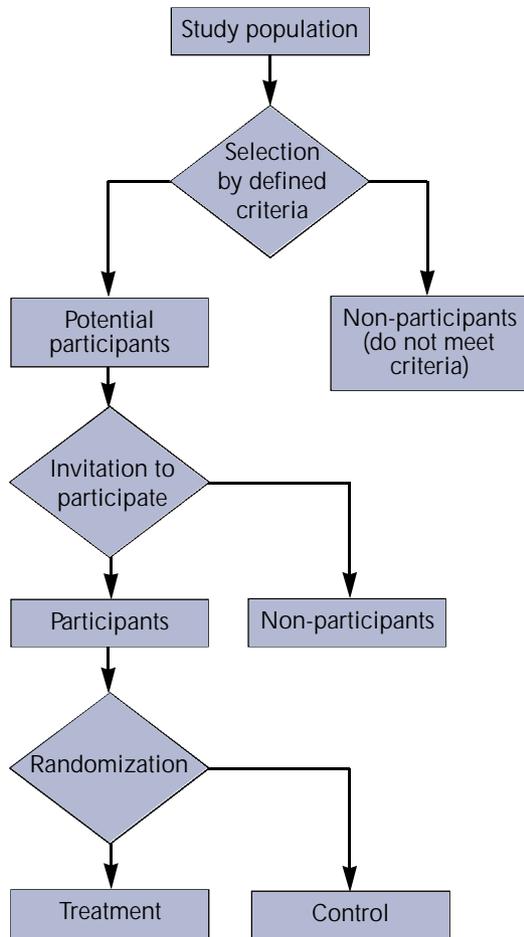
Although the findings of the two epidemiological studies are impressive, especially when considered in conjunction with the biochemical evidence, they do not constitute definitive proof that vitamin E protects against heart disease. The two studies were observational in nature; they were not randomized intervention trials. The participants in the studies made their own decisions about whether or not to use vitamin E supplements. It is conceivable that people who chose to take supplements might have differed from the others in ways that influence heart disease risk. These other factors, rather than vitamin use itself, could have been responsible for the observed associations. Although the researchers made extensive efforts to measure and control for the effects of many dietary and lifestyle factors that may influence heart disease risk, it remains possible that some unmeasured factor may have biased the results.

In addition, the two epidemiological studies relied exclusively on questionnaires as a source of exposure data. Physical examinations of the subjects were not performed, and blood and tissue samples were not collected. For this reason, no one knows whether the participants who reported taking vitamin E supplements actually had higher blood or tissue vitamin E levels than the others. No one knows whether the susceptibility of LDL to oxidation in these individuals was any different from that in non-users of supplements. The full chain of events that might link vitamin E supplementation with a reduced risk of heart disease was not demonstrated in these studies.

A definitive answer to the vitamin E–heart disease question can come only from randomized intervention trials. Two large-scale trials of vitamin E supplementation have already been completed, and neither showed a protective effect against heart disease. The doses of vitamin E used in those trials, however, were relatively low – 30 and 50 mg/day, respectively. If protective effects of vitamin E are seen only at doses of hundreds of milligrams per day, as both the cohort studies and the biochemical experiments suggest, the negative results of these trials are not really relevant. Definitive conclusions must await the completion of additional trials using higher doses of vitamin E. One such trial, involving more than 40 000 United States women, is already in progress.

FIGURE 6

## Design of a randomized intervention trial



Source: Beaglehole R, Bonita R, Kjellström T, Basic Epidemiology, (© Geneva, World Health Organization, 1993).

For example, after British research suggested that children who had received an intramuscular injection of vitamin K at birth had an increased risk of childhood cancer, a United States research team conducted a nested case-control study among the 54 000 subjects in an earlier study of infancy and childhood. Data on vitamin K exposure were obtained from the extensive medical records collected during the original study and the vitamin K exposure of 48 subjects who developed childhood cancer (cases) was compared with that of 240 controls randomly selected from the cohort. No association was found between vitamin K injections and childhood cancer.

## Intervention trials

Intervention trials differ from all of the types of studies discussed thus far in this monograph in that they are experimental in nature; the researchers recruit subjects and then randomly assign them to receive or not receive a treatment under investigation. Figure 6 shows the plan for an intervention trial. Ideally, the participants in a trial should not know their group assignment; this is accomplished by giving an identical-appearing but inactive placebo to those who do not receive the active treatment and by blinding the subjects to their assignment (that is, not telling them until after the study ends). After a period of time, those who received the treatment are compared with those who did not, to see if their health outcomes were different.

## Community trials

A few intervention trials have been carried out at the community level. The trials of water fluoridation are a classic example. In these trials, entire communities had fluoride added to their water supplies, while other communities received untreated water. After a period of time, dental health was assessed in both groups of communities and fluoride was found to be effective in reducing the risk of dental caries.

### Individual-based trials

More commonly, however, subjects are assigned individually to intervention or control groups. For example, researchers in Hungary conducted a trial in which thousands of healthy women who were planning to become pregnant were randomly assigned to receive a daily multivitamin containing folic acid or an identical-appearing placebo containing no vitamins. After more than 4 000 of the participants became pregnant, the researchers found that six women in the placebo group had conceived children with neural tube defects, compared with zero women in the multivitamin group, a statistically significant difference.

### Advantages of intervention trials

The key advantage of intervention trials is that they can provide direct evidence of a cause-and-effect relationship. If the subjects are randomly assigned to the treated and control groups and if blinding is successful, one can assume that any difference that develops between the members of the two groups is directly attributable to the factor under investigation.

### Disadvantages of intervention trials

Unfortunately, however, intervention trials have a few disadvantages, especially when applied to certain types of nutritional factors. One problem is that subjects cannot be blinded to some types of nutritional interventions. For example, in the Oslo study, men were randomly assigned to an experimental group that received individual counselling about smoking cessation and heart-healthy diets or to a control group that received no special counselling. It is impossible to administer an intervention of this type in a blinded fashion. (Despite this and despite the fact that some of the controls stopped smoking or modified their diets on their own initiative, the incidence of heart attacks in the treated group was 47% lower than in the control group.)

Some interventions can be administered in a blinded fashion by putting the active agent into capsules, but this

approach may be less than satisfactory if the real focus of interest is foods. Supplements of  $\beta$ -carotene, calcium or fish oil may be convenient to administer, but their effects may not be the same as those of the vegetables, dairy products or seafoods of real interest to nutrition researchers.

Intervention trials involve ethical considerations that do not apply to observational studies. Unless the experimenters have good reason to believe that their intervention may be helpful, it is unethical to impose it on people. At the same time, the experimenters must be less than certain that the intervention is effective or it would not be ethical to withhold it from the controls. For example, even though scientists might like to conduct additional trials of folic acid supplementation around the period of conception, it is extremely unlikely that any ethics committee would approve such studies. The beneficial effect of folic acid has been established so clearly that it would be unethical to deny this treatment to control subjects.

In many instances, dietary factors exert their effects over a prolonged period of time. It may be necessary to continue an intervention for many years to detect an effect. For example, in a trial in Finland designed to evaluate the effect of  $\beta$ -carotene supplementation on lung cancer risk in male smokers, the subjects received  $\beta$ -carotene or a placebo for an average of 6 years. (Unexpectedly, those who received  $\beta$ -carotene showed a small but statistically significant increase in lung cancer risk – a finding that has not yet been satisfactorily explained.)

A final limitation of intervention trials is that they can assess only one or two factors at a time. Because of limited resources, only a few of the many intriguing hypotheses about diet and disease can be tested in this way. In the remaining instances, scientists must make decisions about probable causality on the basis of observational data, combined with evidence from animal and biochemical studies.

### Sample size considerations

For all of the study designs described above, it is crucial for researchers to choose a sample of adequate size to allow effects to be detected. Researchers must use statistical techniques in combination with their knowledge of the anticipated strength of the association under investigation to determine the correct number. If too many subjects are included, the study may become inordinately expensive; if too few, the study will be worthless.

## Exposure assessment and data quality

### Dietary intake

In all types of epidemiology, scientists need to measure both exposures and outcomes. In nutritional epidemiology, assessment of exposure poses particular problems because dietary intakes are complex and must be measured in special ways.

Nutritional exposures may be defined in different ways for different research purposes. Depending on the focus of the study, the exposures may be the foods that people eat, the nutrient or non-nutrient components of those foods, anthropometric measures such as height and weight, biochemical measures of nutritional status or clinical assessments. Epidemiologists need to define their research questions clearly to choose the most appropriate exposures.

Because food consists of many substances, not just nutrients, it is not adequate to equate nutrient intake with food intake. For example, even though fruits are the most important sources of vitamin C in most people's diets, it would be incorrect to equate vitamin C intake with fruit intake. Associations between fruit intake and disease risk might be due to components of fruit other than vitamin C. Similar distinctions need to be made with regard to non-nutritive components of food. For example, coffee intake should not be equated

with caffeine intake, since coffee contains a wide variety of chemical substances other than caffeine and since total caffeine intake may come from several sources, not just coffee.

### Food consumption data

In ecological studies, researchers sometimes rely on national food consumption or food "disappearance" statistics as a measure of dietary intake. These statistics, however, are crude. Their quality varies from country to country and they reflect waste and spoilage of food as well as actual consumption.

For individual-based studies, researchers must use more accurate methods of assessing dietary intake. The methods in current use include dietary recalls, food records, dietary histories and food frequency questionnaires.

### Dietary recalls

For a dietary recall, subjects are asked to list the foods they consumed during a certain period of time, usually the preceding 24 hours. This method is relatively quick and simple and does not require the researcher to have prior knowledge of the subjects' food habits. Underreporting of food intake may be a problem, however. Increasing evidence indicates that many people systematically underreport their total food intakes and that this tendency is stronger in some segments of the population (especially overweight people) than in others. Dietary recalls are best suited to obtaining information on present diet rather than diet in the distant past, when lapses of memory could be a problem. A single 24-hour recall is not adequate for measurement of an individual's usual intake; multiple recalls must be used. A single 24-hour recall from each individual can be used, however, to estimate the mean nutrient intakes of groups of people rather than of specific individuals.

### Food records

For a food record, subjects record their intake as they eat. They may be asked to estimate portion sizes, to weigh food before they consume it or even to provide a duplicate meal for later analysis. If records are obtained for a sufficient number of days and if subjects cooperate well, food records can provide a good picture of usual current dietary intake. In fact, scientists often consider the weighed food record to be the “gold standard” for measuring food intake. However, this method imposes considerable burdens on the subject and can be used only with relatively educated, literate people. Also, subjects may not comply with the researchers' instructions to maintain their customary eating habits during the study period and some subjects may underreport their intakes.

### Diet histories

For a diet history, subjects are asked open-ended questions regarding their usual (present or past) dietary intake. The interviewer carefully inquires about food consumption meal by meal, seeking information about variations in intake and trying to establish the usual pattern of consumption. Interviews of this type are time-consuming, often lasting for more than 1 hour, but they can provide considerable detail about an individual's eating habits, including subtle aspects such as food preparation practices and seasonal variations in food choices. However, dietary histories, unlike records or recalls, require subjects to make judgments about their usual food habits. Thus, the answers may reflect what the subjects think they eat (or what they would like the interviewer to think they eat), rather than what is actually eaten.

### Food frequency questionnaires

Like diet histories, food frequency questionnaires focus on usual intake. This method is far more structured, however. Subjects complete an interviewer-

administered or self-administered questionnaire that asks how frequently they consume a series of foods. Some types of food frequency questionnaires pose their questions in open-ended form, whereas others use closed-ended questions with predetermined response categories. For example, in the open-ended type of questionnaire, subjects would be asked how often they eat apples. In the closed-ended type of questionnaire, they might be asked whether they eat apples daily, four to six times per week, one to three times per week, less than once per week but more than once per month, less than once per month or never.

Food frequency questionnaires provide a reasonable measure of usual present or past intake. They are commonly used in epidemiological research because it is easy for the researcher (or a computer) to transform the answers into usable data. The quality of the results, however, depends greatly on the quality of the questionnaire.

To design a good questionnaire, epidemiologists need considerable advance knowledge of their study subjects. If the study population is to include immigrants, for instance, the researchers will need information about their countries of origin so that they can include appropriate ethnic dishes in the questionnaire. Researchers also need an accurate knowledge of the food sources of the nutrients that they wish to study so that they can include the appropriate foods in the questionnaire. For example, it has been reported that some questionnaires currently used in the United States to assess carotenoid intake may yield misleading results because they omit certain mixed dishes (for example, vegetable soup) that contribute substantially to total carotenoid intake.

### Nutrient databases

If researchers wish to analyze their findings in terms of nutrients, they need to convert the food consumption

data from their recalls or questionnaires into nutrient intake data. This requires the use of food composition tables or nutrient databases, which can be obtained from a variety of sources.

Unfortunately, the quality of food composition data varies from nutrient to nutrient. For example, until recently, data on the amount of vitamin E in foods were so inadequate that few researchers attempted to assess dietary intakes of this vitamin. Even for nutrients for which extensive data are available, the values in databases may need to be refined as new scientific information is developed. For example, recent analyses using high-performance liquid chromatography, the state-of-the-art method of analysis for vitamin C, have produced values lower than those found in existing databases, indicating that the use of the databases may lead to overestimation of vitamin C intake.

Of course, if researchers express their findings in terms of foods rather than nutrients, they need not be concerned about the potential inadequacies of nutrient databases. For example, the United States researchers who conducted a case-control study of cataract did not need to use a nutrient database to determine that people who ate fewer than two servings of vegetables per day were at increased risk of this eye disease. Expressing findings in terms of foods also helps researchers avoid making unwarranted assumptions about exactly what the active agent might be, and it makes it easy to translate findings into dietary recommendations. Even though the components of vegetables that protect against cataracts have not been identified, the results of the case-control study are of practical value because they add to the evidence supporting recommendations that people should eat several servings of vegetables daily.

### Biomarkers

With all the difficulties involved in measuring dietary intake, it might seem simpler to assess nutritional status

using biomarkers, such as blood or urine levels of a nutrient. Indeed, biomarkers are often used in epidemiological studies. In some instances, they may even be the preferred way to evaluate exposure to a particular dietary factor. Sodium is a good example. Because it is difficult to determine the sodium content of the diet accurately, several important studies, including the Intersalt study (an enormous study of sodium and blood pressure involving more than 10 000 subjects in 32 countries), have used urinary sodium excretion as their measure of sodium intake.

Unfortunately, the use of biomarkers is almost as complex as the assessment of diet. A wide variety of factors influence the accuracy and interpretation of these indicators of dietary exposure.

Most of the currently available biomarkers reflect nutrient intakes over a relatively short period of time. For example, the levels of certain nutrients in urine or blood serum may change within hours or days of a change in dietary intake. Other biomarkers, such as the levels of nutrients in red blood cells, change more slowly, reflecting dietary intake over a period of weeks or months. Unfortunately, few of the currently available biomarkers reflect truly long-term intakes over a period of years or decades. This is an important limitation, since long-term patterns of dietary intake are of great interest in studies of chronic diseases.

Some biomarkers are good measures of recent exposure. Urinary levels of potassium, sodium, fluoride, chloride and iodide are considered good indicators of the intakes of these minerals. Other biomarkers, however, are of little value. For instance, except in deficiency situations, blood vitamin A levels are unrelated to dietary vitamin A intake. For some exposures of interest to nutritional epidemiologists (for example, calcium, magnesium), no biochemical indicators have been established.

Some theoretically valuable biomarkers may be of little use because their collection is unacceptable to study subjects. The concentrations of nutrients in the liver are of great interest, for example, but they are rarely measured in healthy individuals because few people would be willing to have a sample of their liver taken solely for research purposes.

Some indicators are influenced by factors other than dietary intake. For example, smoking reduces blood levels of vitamin C and carotenoids. In such instances, the biomarker may still be a good indicator of nutritional status, but it may not accurately reflect dietary intake.

Contamination of samples may occur even before the specimens are collected and may occur after specimen collection unless appropriate precautions are taken. For example, specially prepared tubes must be used in the collection of blood for zinc measurement because standard tubes may be contaminated with zinc.

Some biomarkers, such as vitamin C, are easily degraded during storage. Unlike many other nutrients, vitamin C cannot be measured in blood samples that have been subjected to prolonged frozen storage.

Finally, biochemical indicators of dietary intake are only as good as the laboratories that measure them. A quality assurance programme is an essential part of any programme of laboratory analysis. Researchers involved in multicentre or multinational studies also need to pay attention to whether the assessments performed in different laboratories are comparable. Comparability is so important and so difficult to achieve that some researchers choose to avoid the issue completely by having all analyses performed in the same laboratory, even if this means shipping some samples halfway around the world.

## Data analysis

### Objectives of data analysis

The objectives of data analysis are to determine whether associations exist between exposures and outcomes and to assess the strength of the associations. The direction and strength of an association may be expressed as a relative risk or an odds ratio.

### Relative risk

Relative risk is the ratio of the outcome rate among persons exposed to a certain factor divided by the outcome rate among persons not exposed. If the relative risk is greater than one, people exposed to the factor have an increased risk of the outcome under investigation. If the relative risk is less than one, people exposed to the factor have a decreased risk of the outcome.

For example, in a Spanish study of bladder cancer, subjects with high saturated fat intake had a relative risk of 2.25, meaning they had more than double the risk of developing bladder cancer than did those with low saturated fat intake. In an Italian study of colorectal cancer, subjects with high intakes of  $\beta$ -carotene had a relative risk of 0.38, meaning they had about one-third the risk of developing colorectal cancer compared to those with low  $\beta$ -carotene intakes.

Relative risks can be used to compare the strength of different associations. The relative risk of lung cancer in people with low fruit and vegetable intake compared to those with high intake is about 2.0. The relative risk of lung cancer in smokers compared to non-smokers is at least 10.0. Clearly, the association with smoking is stronger than the association with fruit and vegetable intake.

From a public health policy standpoint, it makes sense to focus preventive efforts on the most important risk factors for a particular disease. For example, in the case

of lung cancer, prevention should focus on the elimination of cigarette smoking, since this factor has the greatest impact on the relative risk of this disease. Encouraging fruit and vegetable consumption would also be helpful, but it is less important than smoking cessation.

### Odds ratios

Instead of relative risk, some case-control studies use the odds ratio, which is the ratio of the odds of exposure for cases to the equivalent odds for controls. For relatively rare diseases, the odds ratio is a good estimate of the relative risk. As with relative risk, odds ratios greater than one indicate an increase in risk and odds ratios less than one indicate a decrease in risk.

For example, in one recent study, women with low blood levels of  $\beta$ -carotene had an odds ratio for cervical cancer of 3.1, meaning that their likelihood of cervical cancer was increased. In another study, subjects who had a history of being breast-fed in infancy had an odds ratio for breast cancer of 0.74, indicating that their likelihood of breast cancer was decreased. Because the incidence rates of cervical and breast cancers in the study populations were low, these odds ratios can be used as estimates of relative risks.

### Attributable risk

Another measure of the relationship between exposure and disease, with a meaning different from that of the relative risk or odds ratio, is attributable risk – the extent to which the occurrence of a disease or other outcome variable can be attributed to a particular factor. A related measure, the population attributable risk proportion, indicates the proportion of all cases in a total defined population that can be attributed to the factor. This measure reflects both relative risk and the frequency of the factor in the population and it is useful for determining whether efforts to modify a factor would be likely to have a substantial impact on public health.

For example, in a study of laryngeal cancer in Italian men, attributable risks were estimated for three important risk factors: smoking, alcohol and dietary  $\beta$ -carotene intake. Smoking accounted for 77% of total risk, alcohol for 25%, low  $\beta$ -carotene intake for 18% and the three factors combined for 86% of all laryngeal cancers. If the relationships between these risk factors and the disease prove to be causal, efforts to modify these three factors (especially smoking) might substantially reduce the risk of laryngeal cancer in the study population.

### Confounding

To obtain valid results from data analysis, it is necessary to consider the possible effects of confounding factors (factors that are unequally distributed in the groups under study and that may give rise to spurious associations). Researchers may deal with this problem by analyzing data separately for subjects who fall into different categories (strata) in terms of the confounding factor. For example, epidemiologists often analyze data separately for men and women, for smokers and nonsmokers or for different age groups.

In many situations, researchers use statistical techniques to adjust for the effects of confounding factors during analysis. For instance, most epidemiologists who are analyzing a study of diet and breast cancer would want to adjust for reproductive factors known to influence breast cancer risk, such as age at menarche, age at first childbirth and parity (number of births).

Energy intake is an important confounder in many nutritional epidemiology studies. Total energy intake is positively correlated with intakes of most nutrients. The interpretation of epidemiological findings may change depending on whether this fact is taken into account. In a recent trial, researchers fed women varying numbers of eggs (or a placebo consisting of cholesterol-free egg substitute) and measured the effect on serum cholesterol levels. They found that serum cholesterol increased by

2.81 mg/dl for each 100 mg of egg cholesterol added to the women's diets – a much greater increase than the 1.47 mg/dl observed in a similar study of men. It might appear from these values that women are more responsive than men to the cholesterol-raising effects of eggs, but this interpretation is probably not correct. The women had lower energy intakes than the men and therefore the added eggs represented a larger proportion of their total food consumption. To make a meaningful comparison between the results for the two sexes, it would be necessary to adjust for the difference in energy intake.

Epidemiologists must be careful not to treat a factor as a confounder when it is actually a part of a causal pathway. For instance, one of the ways in which obesity may increase the risk of heart disease is by causing an increase in blood pressure, so it would not be desirable to adjust for blood pressure in a study of obesity and heart disease.

### Steps in data analysis

To analyze their data in a meaningful way, epidemiologists must use strategies that enable them to account for the effects of known confounding factors and perhaps to discover some others, as well as to estimate the effects of the factors under investigation. The possibility of interactions among different factors must also be considered.

Multivariate analysis is often used in situations where several variables must be accounted for simultaneously. The techniques of multivariate analysis differ, ranging from simple cross-classification and adjustment to more complex methods of statistical regression analysis. Multivariate techniques allow researchers to determine which of the variables has an independent association with the outcome, to detect interactions among variables and to measure the relative contribution of each variable to the risk of the disease. It is important to note, however, that multivariate analysis, like univariate

analysis, does not distinguish causal from noncausal associations. Also, dealing with interactions is difficult. If interactions are present, it may be better to present stratified results for different segments of the study population (for example, smokers versus non-smokers).

## INTERPRETATION

---

When interpreting the findings of epidemiological studies, scientists must consider two types of validity: internal validity, or the degree to which the findings are true for the study subjects and external validity, or the degree to which the findings can be extended to persons other than the study subjects.

### Internal validity

Many questions must be asked to assess the internal validity of a study. Have the exposure and outcome variables been measured accurately? Is it likely that selection bias or information bias may have occurred? Have relevant confounders been identified and measured accurately? Has proper adjustment been made for the effects of confounders? If measurement error or bias might have occurred or if an unmeasured confounder might have been present, how would this affect the results? In what direction might the estimates be biased?

If bias might merely have weakened the strength of an association, its effect might not be crucial to the interpretation of the study's findings. If bias might have strengthened an association or created a false association, however, it might invalidate the study's results. The following two examples illustrate these contrasting situations.

A case-control study demonstrated a significant negative association between dietary folate intake and the risk of heart attack in a group of men and women

from the eastern United States (This association, which is most likely mediated by the effect of folate nutrition on blood levels of the amino acid homocysteine, has also been observed in several other studies.) There is good reason to question the accuracy of the assessment of dietary folate intake in this study. Folate is very labile; seemingly small differences in the processing, cooking or storage of foods can lead to large differences in folate content. Because of the great variability in the folate content of foods, there might have been some non-differential (random) misclassification of the folate intakes of the subjects. However, this misclassification would not invalidate the study's results, because it would tend to weaken, rather than strengthen, any association between folate and heart disease. If anything, the true association might be even stronger than that reported in this study.

A large case-control study conducted by the United States National Cancer Institute showed that people who regularly took vitamin E supplements had a substantially and significantly lower risk of oral cancer than those who did not take them (see Figure 7). Does this mean that vitamin E protects against oral cancer? Not necessarily. People who choose to take supplements are generally more health conscious than those who do not; their lifestyles tend to be more healthful in many ways, only some of which can be measured and accounted for in an epidemiological study. If other health practices are responsible for some or all of the reduction in cancer risk observed in the supplement users, the true effect of vitamin E supplementation would be weaker than that reported in this study. Indeed, there might be no true association at all.

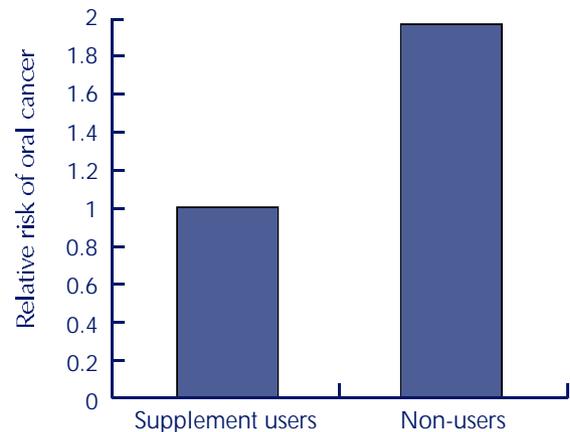
These two examples illustrate the general point that non-differential misclassification (random error that is equally likely to occur in all subgroups of study subjects) is of less concern than differential misclassification (non-random error that affects some subgroups of study subjects more than others).

Differential misclassification can result in a wide variety of erroneous conclusions because it can either strengthen or weaken an association and it can even create a spurious association where none really exists. In almost all circumstances, however, non-differential misclassification can only weaken an association.

To appreciate the limited effect of non-differential misclassification, it may help to imagine that you are attempting to draw a picture of a horse while travelling in an automobile over a very bumpy road. The random jostling of your hand will probably cause you to make errors in your drawing – errors that decrease the resemblance of your drawing to a real horse. This is

FIGURE 7

Oral cancer risk in users and non-users of vitamin E supplements



Source: Gridley G, McLaughlin JK, Block G, et al., Vitamin supplement use and reduced risk of oral and pharyngeal cancer, *American Journal of Epidemiology* 135:1083–1092, 1992.

similar to the weakening of an association by random error in an epidemiological study. The movements of the vehicle might even distort your picture so much that the animal would be unrecognizable, which is similar to the weakening of an association to the point where it is no longer detectable. It is very unlikely, however, that the bouncing of the car would cause you to draw the horse more accurately than you would have done if you were sitting at a desk. Similarly, it's very unlikely that non-differential misclassification in an epidemiological study would strengthen an association. And just as the random jostling of the automobile would not cause you to draw an elephant instead of a horse, non-differential misclassification almost never creates spurious associations.

#### The role of chance

Epidemiologists must also consider the possibility that an observed association between an exposure and an outcome might be due to chance. This possibility is customarily assessed using formal tests of statistical significance. Most often, a finding is considered significant if there is only a 5% (that is, one in 20) likelihood that a result as extreme as that observed would have occurred by chance, in the absence of a true association between the exposure and the outcome.

For example, in a Korean case-control study of stomach cancer, the use of a home refrigerator was significantly associated with decreased risk of the disease. This means that the probability that the observed association was due to chance alone was less than one in 20. (Home refrigeration might reduce risk because it helps to preserve fresh foods, which might provide protective nutrients. The availability of a refrigerator might also reduce dependence on traditional preservation methods such as salting, smoking and pickling, which have been associated with increased stomach cancer risk.)

#### Confidence intervals

When findings are expressed as relative risks or odds ratios, the role of chance is usually evaluated using a confidence interval. The confidence interval represents the range within which the variable is likely to lie. Most commonly, researchers use a 95% confidence interval, which indicates that there is a 95% probability that the true value of the relative risk (or odds ratio) falls between the two stated values. To put it another way, if the study were repeated on many different samples from the same population and the confidence intervals were calculated for each, 95% of these would include the relative risk (or odds ratio) estimated in the first study. For example, a study of United States men evaluated the effects of multivitamin use on the later development of cataracts. Among current smokers, the relative risk of cataract in multivitamin users, compared with nonusers, was 0.38, with a 95% confidence interval of 0.16–0.92. This means that there is a 95% probability that the value of the relative risk lies within this interval.

#### Statistical significance versus clinical significance

Even if an association is statistically significant, it might not be biologically or clinically significant. Some effects are too small or too rare to be of practical importance. In an analysis of data on more than 1 800 participants in a long-term trial of vitamin A supplementation, subjects who received supplements showed a statistically significant increase in serum triglyceride levels compared with those receiving placebo. Because this was an experimental study, it is almost certain that the relationship between the exposure (vitamin A) and the outcome (increased triglycerides) was causal (although triglyceride level was probably not a previously defined endpoint of the study). However, the magnitude of the effect was very small – too small to cause a meaningful change in cardiovascular risk. The investigators concluded that this minor effect was not clinically significant even though it was statistically significant.

Conversely, it is possible for an effect to be of potential clinical significance even if it is not statistically significant. Sometimes an apparent effect may be important enough to justify action, even though no one can be sure whether the effect is real. For example, in early 1996, a United States research team discontinued a randomized trial of  $\beta$ -carotene/vitamin A supplementation when preliminary results showed higher numbers of lung cancers and deaths in the supplemented group than in the placebo group. The increases were not statistically significant and it was entirely possible that they were due to chance. However, the researchers were concerned about these increases because they suggested the possibility of harm to the study participants and because they resembled effects seen in a previous  $\beta$ -carotene trial in Finland. For these reasons, a decision was made to stop the trial, even though it was uncertain whether the supplements had truly increased lung cancer incidence or mortality.

### Power

The term power refers to the relative frequency with which a specified effect would be detected by a proposed study if the study were to be repeated using the same conditions. To be reasonably sure that such an association will be detected, epidemiologists must plan their studies so that they have adequate power, which is largely a matter of choosing a sufficient sample size to give a high probability that true effects will be found. For example, for a trial of vitamin A supplementation in patients with the hereditary eye disease retinitis pigmentosa, the researchers chose a sample size of 601 subjects to give the study a power of greater than 95%. This means that the likelihood was more than 95% that an effect of predefined amplitude would be detected if it were present. Knowing the power of a study is particularly important in the interpretation of negative findings, that is, situations where no association is found. (In this particular trial, however, the results were positive. Vitamin A supplementation significantly slowed the progression of retinitis pigmentosa.)

The power of a study also plays a role in determining whether the confidence interval around a particular value is narrow or wide. If the power of a study is low, the confidence interval is likely to be wide, making it difficult to determine the strength of an association. For example, a small United States study (which included only 61 subjects) found that women with high dietary intakes of animal fat had a significantly higher risk of bone fractures than those with lower intakes. The odds ratio was 5.00, suggesting that the increase in risk was substantial. However, the 95% confidence limits around this value were 1.33–18.82, a confidence interval so wide that it is impossible to know whether the true effect of animal fat was relatively small or extremely large.

### Exposure range and dose-response relation

In some instances, a study may fail to detect an association because intakes of the dietary factor under investigation are too low, too uniform or not in the range where effects are likely. For example, although most epidemiological studies of oral and pharyngeal cancers have found significant negative associations with fruit and vegetable intakes, it is not surprising that a case-control study in Puerto Rico failed to do so. In this population, intakes of fruits and vegetables were extremely low: 75% of the subjects said that they never ate fruit and 87% ate no more than two servings of vegetables per week. It is likely that very few people in this population ate enough of these foods to influence their cancer risk substantially.

The several studies that have examined the association between fish intake and heart disease risk provide a different type of example. In this instance, significant associations have been demonstrated only in populations in which some people eat fish and others do not. In populations where everyone eats at least some fish, no significant associations between fish intake and heart disease risk have been detected. This is interpreted as meaning that eating a small amount of fish is associated with some protection against heart

disease, but that eating a lot of fish is not associated with any additional benefit.

## External validity

The term external validity refers to the applicability of epidemiological findings to people other than the subjects who were studied. To assess external validity, scientists evaluate a study's findings in conjunction with other available evidence. They also consider possible differences between the study subjects and the target population to whom they wish to generalize the results.

When many studies in different populations yield similar results, the findings are likely to be externally valid. The inverse association between fruit and vegetable intake and cancer risk is a good example of this. This association has been demonstrated in more than 130 epidemiological studies conducted in a wide variety of populations from at least 17 nations, including both industrialized and developing societies. The evidence for this association is exceptionally consistent, particularly for epithelial cancers (for example, cancers of the lung, mouth, larynx, oesophagus and stomach). Most nutrition scientists and epidemiologists would agree that this association can be extrapolated with confidence to populations throughout the world.

## Meta-analysis

The interpretation of epidemiological findings is far more difficult when the results of different studies conflict with one another. In this situation, the inconsistency can sometimes be resolved through a meta-analysis – a quantitative technique in which the statistical results of separate studies are pooled to yield overall conclusions. One recent meta-analysis evaluated 15 epidemiological studies that had evaluated the relationship between blood lead levels and systolic

blood pressure in men. Although the original studies did not have consistent results, the meta-analysis showed a significant positive association between lead and blood pressure. The researcher who conducted the analysis concluded, therefore, that it would be prudent to continue efforts to minimize exposure to lead from food, water and other sources.

Meta-analysis may be more objective than traditional critical reviews of the literature and it can help make sense out of studies too small to provide reliable answers when analyzed individually. However, decisions about which studies to include in a meta-analysis can be difficult. Opinions differ on whether flawed studies and unpublished studies (often the ones that give negative results or no effect) should be included in meta-analyses and on whether studies of better quality should be given greater weight than those of lesser quality.

Studies that are to be combined in a meta-analysis should be similar in terms of the types and amounts of exposures and the types of outcomes assessed. If the studies differ in these respects, the meta-analysis may yield erroneous results. For example, a meta-analysis that supposedly indicated that vitamin C has no effect on the common cold has been questioned because it included studies that employed low doses of vitamin C (200 mg/day or less) as well as those that used megadoses (1–5 g/day). The negative results of the low-dose studies may have diluted the more positive results of the high-dose studies, leading to an incorrect conclusion. Critics have argued that it would have been better to analyze the two groups of studies separately.

## Extrapolation to different populations

Epidemiologists and public health authorities should use caution when extrapolating findings from one group of people to other, very different population groups. For example, supplementation with a

combination of  $\beta$ -carotene, vitamin E and selenium caused a significant reduction in cancer mortality in the participants in an intervention trial in Linxian, China, but it would be wrong to conclude on the basis of this trial that people in Western Europe would necessarily reduce their risk of cancer by taking a similar

supplement. Linxian is a poor rural area with unusual patterns of cancer risk and a high prevalence of marginal nutritional deficiencies. Findings obtained in this atypical and nutritionally deprived population may not be generalizable to better-nourished populations with very different patterns of disease.

## BOX 3

### Evaluating the quality of epidemiological studies

Epidemiological findings are only as good as the studies that produce them. To assess the quality of epidemiological studies, readers should ask many critical questions about their design and execution, including the following:

#### Intervention trials

- Were the subjects assigned randomly to the treatment and control groups?
- Was randomization successful (that is, were the treatment and control groups truly comparable with respect to important variables)?
- Were efforts made to determine whether subjects complied with the treatment protocol?
- Was compliance good?
- Were subjects kept "blind" to their treatment assignment through the use of a placebo or other means?
- Was blinding successful (that is, were subjects unable to guess their group assignment correctly)?
- Were the researchers kept "blind" as to the subjects' treatment assignment?
- Was the assessment of outcome blinded?
- Was the number of subjects large enough to yield statistically reliable results?

#### Cohort studies and nested case-control studies<sup>1</sup>

- Was the number of subjects large enough to allow effects to be detected?
- Was the follow-up period long enough to yield meaningful results?
- Did the investigators clearly define the reference population from which the cohort was drawn?
- Were exposures and outcomes assessed by accurate and appropriate methods?
- Were exposure assessments repeated during the study to account for possible changes over time?
- Were many subjects lost to follow-up? If so, could this have affected the study outcome?
- Were data collected on potential confounding variables? Were these variables taken into account in the data analysis?

## BOX 3 – continued

- Did sufficient time elapse between the initial examination and the diagnosis of disease to ensure that metabolic effects of disease could not have influenced biological variables measured in the study?
- How were the diagnoses confirmed in the subjects who developed the disease under investigation?
- If biological samples were collected at the outset of the study and stored for later analysis, were suitable storage conditions used? Were all samples handled in the same manner?

### Case-control studies<sup>1</sup>

- Was the number of subjects large enough to allow effects to be detected?
- How was the diagnosis confirmed in the cases?
- Were exposures and outcomes assessed by accurate and appropriate methods?
- Were the rates of participation in the study (by both potential cases and potential controls) high? If not, could this have influenced the study outcome?
- Were controls selected from the same population that yielded the cases?
- Were controls comparable to cases in all respects other than those under investigation?
- Was information collected in the same way from cases and controls?
- Did the subjects themselves provide the exposure data, or was it necessary to resort to proxy respondents?
- Did the assessment cover an appropriate time frame?
- Were interviewers blinded to the subjects' case or control status?
- Did the researchers consider the possibility that metabolic effects of disease could have influenced measurements such as blood pressure or blood nutrient levels?
- Was the purpose of the study defined beforehand rather than after the data were collected?

### Studies (of any design) that showed a statistical association between an exposure and an outcome

- Could the association have been due to chance?
- Could the association have been due to bias?
- Could the association have been due to confounding?
- Is the association biologically plausible?
- To whom does the association apply?
- Is the association likely to represent a cause-and-effect relationship? (See Table 2 for a further discussion of criteria for assessing the likelihood of causality.)

### Studies (of any design) that did not show a statistical association between an exposure and an outcome<sup>2</sup>

- Did the study have the power to detect a clinically or biologically significant effect if it were present?

## BOX 3 – continued

### Reviews and meta-analyses<sup>3</sup>

- Were the question(s) and methods clearly stated?
- Were comprehensive search methods used to locate relevant studies?
- Were explicit and appropriate methods used to determine which papers to include in the review?
- Was the methodologic quality of the primary studies assessed?
- Were the selection and assessment of the primary studies reproducible and free from bias?
- Were differences in individual study results adequately explained?
- Were the results of the primary studies combined appropriately?
- Were the reviewers' conclusions supported by the data cited?

1. Recently, a group of epidemiologists developed a quantitative scoring system to judge the scientific quality of case-control and cohort studies of nutrition and disease. In this system, case-control studies are scored in three areas (dietary assessment, recruitment of subjects and analysis), and cohort studies are scored in four areas (dietary assessment, definition of cohort, ascertainment and analysis). For more information on this scoring system and its application see Margetts BM, Thompson RL, Key T, et al, Development of a scoring system to judge the scientific quality of information from case-control and cohort studies of nutrition and disease. *Nutrition and Cancer* 24:231–239, 1995.

2. For more information see Friedman GD, *Primer of Epidemiology*, 4th ed. (© New York:McGraw-Hill, 1994), p321.

3. For more information see Sackett DL, Haynes RB, Guyatt GH, Tugwell P, *Clinical Epidemiology: A Basic Science for Clinical Medicine*, 2nd ed. (© Boston: Little, Brown & Company, 1991), p380.

## Causality

In the absence of intervention trials, scientists use several criteria to evaluate whether an observed association between an exposure and an outcome is likely to be causal. These criteria include the strength and consistency of the association, the presence or absence of a dose-response relationship, the temporal relation between exposure and outcome and the biological plausibility of a causal relationship. Table 2 lists some guidelines for assessing whether an association is likely to be causal.

The stronger the association between an exposure and an outcome, the greater the likelihood that it is causal. An association consistently observed in different populations and under different circumstances also suggests that it may be causal. If the intensity of a response increases with dose, it might be more likely that a causal relationship exists. It is important to note, however, that the absence of a dose-response relationship does not prove that an association is not

causal; in some situations, a threshold effect exists and no dose-response relationship would be expected.

If a relationship is truly causal, an outcome should occur only after the exposure. Temporal relationships should also be considered in terms of the underlying aetiology of the disease. For example, cancer is a chronic disease with a long latent period. One would not expect the incidence of cancer to increase within weeks or months after the onset of exposure to a cancer-causing agent.

Finally, a causal association between an exposure and an outcome is more plausible if there is a known or hypothesized biological mechanism by which the exposure is likely to alter the risk of a disease.

The findings of observational epidemiology studies should be considered in conjunction with the results of animal experiments, biochemical studies and other types of research to provide a complete picture of the causation of a disease. Observational epidemiology alone cannot prove that a relationship is causal.

TABLE 2

## Guidelines for causation

|                            |   |
|----------------------------|---|
| Temporal relation          | Does the cause precede the effect?  |
| Plausibility               | Is the association consistent with other knowledge?                           |
| Consistency                | Have similar results been shown in other studies?                             |
| Dose-response relationship | Is increased exposure to the possible cause associated with increased effect? |
| Reversibility              | Does the removal of a possible cause lead to reduction of disease risk?       |
| Study design               | Is the evidence based on a strong study design?                               |
| Judging the evidence       | How many lines of evidence lead to the conclusion?                            |

Source: Beaglehole R, Bonita R, Kjellström T, Basic Epidemiology (© Geneva, World Health Organization, 1993).

## APPLICATION OF NUTRITIONAL EPIDEMIOLOGY DATA

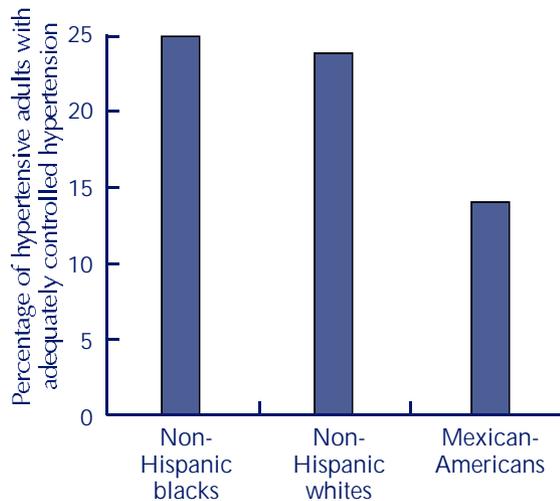
As the examples in this monograph have illustrated, nutritional epidemiology studies can generate information that is of great relevance to public health. The findings must, however, be interpreted with caution. Scientists and policy makers need to appreciate the inherent limits of epidemiology in the detection of weak associations and the complexities involved in measuring dietary intake, avoiding bias, dealing appropriately with confounding factors, analyzing data and assessing causality. If nutritional epidemiology research is misinterpreted by the lay public or by non-epidemiologically trained scientists, it may do more harm than good. However, if nutritional epidemiology is applied in appropriate ways, it can be of great value to public health authorities, health care professionals and the food industry. The three situations described below illustrate ways in which findings from nutritional epidemiology might be applied by these different groups.

### Hypertension in the United States

An analysis of data from the most recent NHANES survey (NHANES III) has indicated that hypertension is far more likely to go untreated or to be inadequately controlled in Mexican-Americans than in other segments of the United States population (see Figure 8). This finding will probably prompt United States agencies to target more of their educational efforts about the importance of blood pressure control toward the Mexican-American community.

FIGURE 8

Percentage of hypertensive United States adults with adequately controlled hypertension



Source: Burt VL, Whelton P, Roccella EJ, et al, Prevalence of hypertension in the U.S. adult population, *Hypertension* 25:305-313, 1995.

The same survey showed that substantial numbers of Americans with hypertension are taking advantage of non-pharmacological methods of treatment, including weight control, dietary sodium restriction and moderation of alcohol intake. Some 3.6% of Americans with hypertension appear to have controlled their blood pressure through these measures alone. These findings will probably prompt health professionals to encourage more of their patients to try non-pharmacological approaches to blood pressure control. They may also prompt the food industry to continue to develop and market reduced-calorie and reduced-sodium versions of their products.

## Folic acid and neural tube defects

A study in Leeds, United Kingdom, indicated that only 19% of a group of 603 women who became pregnant in 1994 had increased their intake of folic acid before conception. A study in South Carolina, United States, showed that only six of 71 women (8.4%) who conceived children with neural tube defects had taken a folic acid supplement during the periconceptional period. Even though these studies were conducted almost 2 years after United Kingdom and United States authorities recommended increased intakes of folic acid for women of childbearing potential, the vast majority of women in the two study populations had not received or acted on this important message. These disturbing findings may increase the impetus for folic acid fortification of staple foods. They may also prompt individual health professionals to make greater efforts to ensure that the women under their care receive information about the need for adequate folic acid intake and guidance on how to achieve it.

## $\beta$ -Carotene

For some years, it was widely believed that the inverse association between fruit and vegetable intake and the risk of various diseases was probably attributable in large part to the  $\beta$ -carotene content of these foods. Animal experiments and biochemical studies had suggested plausible mechanisms of action of  $\beta$ -carotene and dietary or blood levels of  $\beta$ -carotene had been correlated with reduced risks of a variety of diseases in human populations.

More recently, however, some epidemiological studies have cast doubt on this idea. As previously mentioned, not only did large intervention trials in Finland and the United States fail to show a beneficial effect of  $\beta$ -carotene supplementation on lung cancer risk in smokers, but they actually suggested the possibility of a small adverse effect. A third intervention trial in a

## BOX 4

### Nutritional epidemiology in action: iron and heart disease

Does good iron nutriture increase the risk of coronary heart disease? A study from Finland, published in 1992, suggested that it might. In that study, blood samples were taken from 1 931 men and analyzed for serum ferritin, an indicator of iron storage. During three years of follow-up, 51 of the men suffered heart attacks. Men with high serum ferritin levels had a significantly higher risk of heart attack than those with lower serum ferritin levels (relative risk 2.2; 95% confidence interval 1.2–4.0).

This finding was understandably disturbing to nutrition scientists and policy makers. These experts have long recommended that people should include ample amounts of iron-rich foods in their diets to avoid iron deficiency. Now it suddenly appeared that this advice might be harmful to people's health. In fact, an editorial in the journal that published the Finnish study suggested that the new findings might prompt a reexamination of the definition of normal iron nutriture. According to the editorial writer, "Perhaps iron depletion, defined as the absence of iron stores without anemia, should be regarded as physiologically normal iron status."

The editorial writer was not the only one to suggest that substantial changes in dietary habits and public health policy might be warranted. Numerous reports in the news media raised the possibility that recommended allowances for iron, food fortification programmes and other policies might need to be modified. Health authorities did not actually implement any of the suggested changes, however. It would have been premature to make major changes in public health policy on the basis of a single epidemiological observation. Even though the idea of a link between iron and heart disease is biologically plausible (in fact, the existence of such an association had been hypothesized more than a decade earlier), a single, unconfirmed study is not a sufficient basis for action.

In the years since the Finnish study was published, epidemiologists have conducted additional studies of the relationship between iron status and cardiovascular disease, both in Europe and in North America. One U.S. study provided some support for the hypothesis of an association between iron and heart disease, but the pattern observed was different from that seen in Finland. (High intakes of one type of dietary iron – haem iron – were associated with increased risk, but total iron intake was not.) Other studies found no association between iron and cardiovascular risk, and still others suggested that high iron status might even be protective. The total evidence currently available (as of mid-1996) does not support the hypothesis that good iron nutriture promotes heart disease.

If public health authorities had changed their dietary recommendations when the Finnish findings were first published, they would have made an unwise decision. Although the Finnish report seemed alarming at first, scientists and policy makers acted prudently when they chose to wait for further data before deciding whether to take action.

largely nonsmoking population showed no association between  $\beta$ -carotene supplementation and the risk of chronic diseases. A case-control study of macular degeneration (a degenerative disorder of the retina that is a major cause of visual impairment in the elderly) found an inverse association with dietary intakes of lutein and zeaxanthin (two carotenoids found in leafy green vegetables) but not with  $\beta$ -carotene. Similarly, a large cohort study of United States women indicated that risk of cataract was inversely associated with intake of spinach (which is rich in lutein but not  $\beta$ -carotene) but not with intake of carrots (the richest source of  $\beta$ -carotene).

These findings have dampened the previous enthusiasm over the possible health benefits of  $\beta$ -carotene. Many scientists now suspect that  $\beta$ -carotene may be primarily a marker for vegetable and fruit intake and that other carotenoids and non-carotenoid components of these foods may prove to be more important in disease prevention than  $\beta$ -carotene itself.

Public health authorities and health professionals are likely to react to these new developments by re-emphasizing the importance of vegetables and fruits in a balanced diet and by cautioning people that supplements of  $\beta$ -carotene or other single nutrients are not a substitute for consumption of these important foods. Manufacturers of foods and supplements who have contemplated adding  $\beta$ -carotene to their products may now prefer to take a more cautious approach and wait for additional data before making a decision on this issue.

As these situations illustrate, the possibilities for the use of nutritional epidemiology to improve public health are almost endless. Of course, to make the best use of epidemiological data, the limits of this science need to be taken into consideration and the results of epidemiological research should be interpreted with care. If, however, the findings from nutritional

epidemiology are interpreted appropriately and applied judiciously, they can play a major role in today's efforts to improve health through disease prevention.

## SUMMARY

---

Nutritional epidemiology is the study of the nutritional determinants of disease in human populations. It is an exciting branch of epidemiological research because it can provide insight into the causation and prevention of many of today's most crucial health problems, including the chronic diseases of aging. However, it is a very difficult field of research because the exposures of interest – dietary intakes – are extremely complex. Also, many of the associations between dietary factors and disease risks are so subtle that they are difficult to discern using epidemiological techniques.

The goals of nutritional epidemiology include monitoring the food consumption, nutrient intake and nutritional status of a population; generating new hypotheses about diet and disease; producing evidence that supports or refutes existing hypotheses; and assessing the strength of diet-disease associations. Ultimately, the overall goal of nutritional epidemiology is to contribute to the prevention of disease and the improvement of public health, but this goal cannot be reached through the sole use of epidemiology. Most types of epidemiological studies (with the exception of intervention trials) can only identify associations; they cannot prove that an exposure causes a health effect. When intervention trials are not possible, evidence from other lines of scientific research must be combined with epidemiological findings to determine whether an association is causal.

Some of the most common study designs used in nutritional epidemiology include:

- descriptive studies: studies that evaluate the amount and distribution of disease within a population in terms of person, place and time, without the purpose of determining the causes of disease
- ecological studies: studies that evaluate the association between diet and disease using aggregate data on groups of people
- case-control studies: individual-based studies in which researchers identify people who have a disease and otherwise similar people who do not have it and compare their exposures to factors that may have influenced their disease risk
- cohort studies: individual-based studies in which researchers identify a group of people who do not have the disease under investigation, collect information on their exposure to risk factors and monitor them for a period of time to see who develops the disease
- intervention trials: experimental studies in which researchers randomly assign subjects to receive or not receive the nutritional intervention that is under investigation and then observe them for a period of time to see whether the intervention influences the occurrence of disease

All study designs can contribute to the fuller understanding of diet-disease relationships. As a general rule, however, individual-based studies are of greater value than those using aggregate data. Prospective cohort studies are more highly regarded than retrospective case-control studies because the prospective design minimizes the likelihood of bias. However, case-control studies continue to play an important role in nutritional epidemiology because they are less time-consuming and less expensive than cohort studies and because they are applicable to a wider variety of diseases and exposures. Intervention trials are of special interest because they are the only type of epidemiological study that can conclusively establish a causal relationship. Unfortunately, many nutritional variables do not lend themselves to investigation in intervention trials.

Techniques used to measure dietary intakes include dietary recalls, food records, diet histories and food frequency questionnaires. Dietary recalls are appropriate for assessing the intakes of groups of people, but a single 24-hour recall may not give an adequate picture of a specific individual's habitual intake. Food records are often considered the best method of assessing dietary intake, but they impose considerable burdens on the subject and the results may not be accurate if subjects modify their food habits during the time of the study or if they underreport their intakes. Diet histories can provide very detailed information, but they require subjects to make judgments about their usual food habits. Food frequency questionnaires provide less detailed information, but they are well suited for use with large groups of people. These questionnaires must be designed with great care to ensure that all important food sources of the nutrients under investigation are included.

Nutritional status can also be assessed using biomarkers, such as blood or urinary levels of nutrients or their metabolites. However, most of the available biomarkers reflect short-term rather than long-term dietary intake, and this limits their value for some types of research.

The objectives of data analysis in epidemiology are to determine whether associations exist between exposures and outcomes and to assess the strength of the associations. Statistical methods are almost always used in data analysis and care must be taken to consider the possible effects of confounding factors. The technique of multivariate analysis is often used in situations where several variables must be accounted for simultaneously.

When evaluating the validity of epidemiological findings, scientists must consider whether findings may have been due to chance and whether a study had

sufficient power to detect an association if it was present. When the results of epidemiological studies conflict with one another, it may be possible to resolve the inconsistency through meta-analysis – a quantitative technique in which the statistical results of separate studies are pooled in an attempt to yield overall conclusions. Even if a study or a meta-analysis shows a statistically significant association, epidemiologists should use caution when extrapolating findings from one group of people to other, very different population groups.

In the absence of intervention trials, scientists use several criteria to evaluate whether an observed association is likely to be causal. These criteria include the strength and consistency of the association, the presence or absence of a dose-response relationship, the temporal relation between exposure and outcome and the biological plausibility of a causal relationship.

Nutritional epidemiology studies can generate information of great relevance to public health. However, the findings must be interpreted with caution. Scientists and policy makers need to appreciate the inherent limits of epidemiology in the detection of weak associations and the complexities involved in measuring dietary intake, avoiding bias, dealing appropriately with confounding factors, analyzing data and assessing causality. If nutritional epidemiology research is misinterpreted by the lay public or by scientists not trained in epidemiology, it may do more harm than good. However, if it is applied in appropriate ways, nutritional epidemiology can be of great value to public health authorities, health care professionals and the food industry. All of these groups can apply the findings of nutritional epidemiology in ways that can contribute to current efforts to improve health through disease prevention.

## GLOSSARY

---

- Analytical epidemiology:** Epidemiological investigations specifically aimed at studying the determinants of diseases in study populations.
- Bias:** Systematic error resulting in over- or underestimation of the strength of the association between an exposure and an outcome.
- Biomarkers:** Measurements in the human body or its products. Some biomarkers, such as the levels of certain vitamins in blood serum, are used as indices of nutritional status. Others are used as indices of the risk or progression of disease.
- Case-control study:** A study design in which persons with a disease (cases) are compared with those without the disease (controls) to see how their exposures to causative factors may have differed.
- Cohort study:** A study design in which data on exposures to possible risk factors for disease are collected from a group of people who do not have the disease under investigation. The subjects are then followed for a period of time to see whether the later development of disease is related to the factors that were measured.
- Confidence interval:** The range of values within which a variable is likely to lie.
- Confounding factors:** Factors that distort an association because they are associated with an exposure as well as a disease or other outcome.
- Descriptive epidemiology:** The study of variations in the occurrence of disease in terms of person, place and time, without the purpose of establishing causal inference.
- Diet history:** A method of dietary assessment in which subjects are asked open-ended questions about their usual dietary intakes.
- Dietary recall:** A method of dietary assessment in which subjects are asked to recall their food consumption over a specific period of time.
- Ecological study:** A study that compares the rates of exposures and diseases in different populations using aggregate data on exposure and disease, not individual data.
- Epidemiology:** The study of the distribution and determinants of disease in human populations and the application of this study to control health problems.
- External validity:** The generalizability of a study's findings to persons other than the study subjects.
- Food frequency questionnaire:** A method of dietary assessment in which subjects are asked to recall how frequently certain foods were consumed during a specified period of time.
- Food record:** A method of dietary assessment in which subjects record the foods that they consume.
- Incidence:** The number of new cases of a disease during a given period of time in a defined population.
- Internal validity:** The accuracy of a study's findings with regard to the study subjects.
- Intervention trial:** A study in which exposure to the factor under investigation is modified by the investigator; an experimental study.
- Meta-analysis:** A quantitative technique in which the results of several individual studies are pooled to yield overall conclusions.
- Multivariate analysis:** A set of techniques for studying the effects of several factors simultaneously. These techniques range from simple cross-classification and adjustment to more complex methods of statistical regression analysis.
- NHANES:** National Health and Nutrition Examination Survey, United States.

**Odds ratio:** The ratio of the odds of exposure for cases to the equivalent odds for controls. The interpretation of an odds ratio is often similar to that of a relative risk.

**Prevalence:** The number of existing cases of a disease in a defined population at a specified time.

**Relative risk:** The ratio of the outcome rate among persons exposed to a certain factor divided by the outcome rate among persons not exposed.

**Risk:** A general term encompassing a variety of measures of the probability of an outcome. The term risk is usually used in reference to unfavourable outcomes such as illness or death.

**Univariate analysis:** Techniques for studying the effects of a single factor on an outcome variable.

## FURTHER READING

---

A full list of references used to compile this concise monograph is available from ILSI Europe. More detailed information on this subject can be found in the texts listed below.

### General references

Ahlbom A, Norell S. *Introduction to Modern Epidemiology*. Newton Lower Falls, MA: Epidemiology Resources, Inc., 1984

Beaglehole R, Bonita R, Kjellström T. *Basic Epidemiology*. Geneva: World Health Organization, 1993

Freudenheim JL. A review of study designs and methods of dietary assessment in nutritional epidemiology of chronic disease. *Journal of Nutrition* 123:401–405, 1993

Friedman GD. *Primer of Epidemiology*, 4th ed. New York: McGraw-Hill, 1994

Lilienfeld DE, Stolley PD. *Foundations of Epidemiology*, 3rd ed. Oxford: Oxford University Press, 1994

Margetts BM, Nelson M (eds). *Design Concepts in Nutritional Epidemiology*, Oxford: Oxford University Press, 1991

Mausner JS, Kramer S. *Epidemiology – An Introductory Text*, 2nd ed. Philadelphia: WB Saunders Co, 1985

Willett W. *Nutritional Epidemiology*. Oxford: Oxford University Press, 1989

## Specific topics and diseases

Ames BN, Shigenaga MK, Hagen TM. Oxidants, antioxidants and the degenerative diseases of aging. *Proceedings of the National Academy of Sciences USA* 90:7915–7922, 1993

Ascherio A, Rimm EB, Stampfer MJ, et al. Dietary intake of marine n-3 fatty acids, fish intake and the risk of coronary disease among men. *New England Journal of Medicine* 332:977-982, 1995

Block G, Patterson B, Subar A. Fruit, vegetables and cancer prevention: a review of the epidemiological evidence. *Nutrition and Cancer* 18:1-29, 1992

Burt VL, Whelton P, Roccella EJ, et al. Prevalence of hypertension in the United States adult population. *Hypertension* 25:305-313, 1995

Carroll KK. Experimental evidence of dietary factors and hormone-dependent cancers. *Cancer Research* 35:3374–3383, 1975

Czeizel AE. Folic acid in the prevention of neural tube defects. *Journal of Pediatric Gastroenterology and Nutrition* 20:4–16, 1995

Eysenck HJ. Meta-analysis and its problems. *British Medical Journal* 309:789–792, 1994

Gridley G, McLaughlin JK, Block G, et al. Vitamin supplement use and reduced risk of oral and pharyngeal cancer, *American Journal of Epidemiology* 135:1083–1092, 1992

Hemilä H, Herman ZS. Vitamin C and the common cold: a retrospective analysis of Chalmers' review. *Journal of the American College of Nutrition* 14:116–123, 1995

Henderson BE, Ross RK, Pike MC. Toward the primary prevention of cancer, *Science* 254:1131–1137, 1991

Hulley SB, Newman TB. Cholesterol in the elderly: is it important? *Journal of the American Medical Association* 272:1372–1374, 1994

Lichtenstein A. Trans fatty acids, blood lipids and cardiovascular risk: where do we stand? *Nutrition Reviews* 51:340–343, 1993

López-Carrillo L, Hernández Avila M, Dubrow R. Chili pepper consumption and gastric cancer in Mexico: a case-control study. *American Journal of Epidemiology* 139:263–271, 1994

Muir CS, Waterhouse JAH, Mack T, et al (eds). *Cancer Incidence in Five Continents*, vol. 5. Lyons: International Agency for Research on Cancer, 1987

Taubes G. Epidemiology faces its limits. *Science* 269:164–169, 1995

Taylor A. Cataract: relationships between nutrition and oxidation. *Journal of the American College of Nutrition* 12:138–146, 1993